10.7

How Iris Recognition Works

John Daugman
University of Cambridge

Algorithms developed by the author for recognizing persons by their iris patterns have now been tested in many field deployments, producing no false matches in millions of iris comparisons. The recognition principle is the failure of a test of statistical independence on iris phase structure, as encoded by multi-scale quadrature Gabor wavelets. The combinatorial complexity of this phase information across different persons spans about 249 degrees of freedom and generates a discrimination entropy of about 3.2 bits/mm^2 over the iris, enabling real-time decisions about personal identity with extremely high confidence. These high confidence levels are important because they allow very large databases on even a national scale to be searched exhaustively (one-to-many "identification mode"), without making false matches, despite so many chances. Biometrics that lack this property can only survive one-to-one ("verification") or few comparisons. This chapter explains the iris recognition algorithms, and presents results of 9.1 million comparisons among eye images from trials in the United Kingdom, the United States, Japan, and Korea.

1 Introduction

Reliable automatic recognition of persons has long been an attractive goal. As in all pattern recognition problems, the key issue is the relation between interclass and intraclass variability: Objects can be reliably classified only if the variation among different instances of a given class is less than the variation between different classes. For example in face recognition, difficulties arise from the fact that the face is a changeable social organ displaying a variety of expressions, as well as being an active three-dimensional (3D) object whose image varies with viewing angle, pose, illumination, accoutrements, and age [1, 2]. It has been shown that for "mug shot" images taken at least 1 year apart, even the best current algorithms can have error rates of 43% to 50% [14–16]. Against this intraclass (same face) variation, interclass variation is limited because different faces possess the same basic set of features, in the same canonical geometry.

Following the fundamental principle that interclass variation should be larger than intraclass variation, iris patterns offer a powerful alternative approach to reliable visual recognition of persons when imaging can be done at distances of less than a meter, and especially when there is a need to search very large databases without incurring any false matches despite a huge number of possibilities. Although small (11 mm) and sometimes problematic to image, the iris has the great mathematic advantage that its pattern variability among different persons is enormous. In addition, as an internal (yet externally visible) organ of the eye, the iris is well

protected from the environment, and stable over time. As a planar object its image is relatively insensitive to angle of illumination, and changes in viewing angle cause only affine transformations; even the nonaffine pattern distortion caused by pupillary dilation is readily reversible in the image coding stage. Finally, the ease of localizing eyes in faces, and the distinctive annular shape of the iris, facilitate reliable and precise isolation of this feature and the creation of a size-invariant representation.

The iris begins to form in the third month of gestation [13] and the structures creating its pattern are largely complete by the eighth month, although pigment accretion can continue into the first postnatal years. Its complex pattern can contain many distinctive features such as arching ligaments, furrows, ridges, crypts, rings, corona, freckles, and a zigzag collarette, some of which may be seen in Fig. 1. Iris color is determined mainly by the density of melanin pigment [4] in its anterior layer and stroma, with blue irises resulting from an absence of pigment: long wavelength light penetrates while shorter wavelengths are scattered by the stroma. The striated trabecular meshwork of elastic pectinate ligament creates the predominant texture under visible light, whereas in the near infrared (NIR) wavelengths used for unobtrusive imaging at

distances of up to 1 m, deeper and somewhat more slowly modulated stromal features dominate the iris pattern. In NIR wavelengths, even darkly pigmented irises reveal rich and complex features.

The author's algorithms [8–10] for encoding and recognizing iris patterns have been the executable software used in all iris recognition systems so far deployed commercially or in tests, including those by British Telecom, NIST, TSA, Sandia Labs, UK National Physical Lab, Panasonic, LG, Oki, EyeTicket, IrisGuard, Sensar, Sarnoff, IBM, SchipholGroup, Siemens, Byometric, Sagem, IriScan, and Iridian. All testing organizations have reported a false match rate of zero in their tests, some of which involved millions of iris pairings. This chapter explains how the algorithms work, and presents new data on the statistical properties and singularity of iris patterns based on 9.1 million comparisons.

2 Finding an Iris in an Image

To capture the rich details of iris patterns, an imaging system should resolve a minimum of 70 pixels in iris radius. In most deployments of these algorithms to date, the resolved iris radius has typically been 80 to 130 pixels. Monochrome video cameras (480×640) have been used because NIR illumination in the 700-nm to 900-nm band was required for imaging to be unintrusive to humans. Some imaging platforms deployed a wide-angle camera for coarse localization of eyes in faces, to steer the optics of a narrow-angle pan/tilt camera that acquired higher resolution images of eyes. There are many alternative methods for finding and tracking facial features such as the eyes, and this well-researched topic will not be discussed further here. Most images in the present database were acquired without active pan/tilt camera optics, instead exploiting visual feedback via a mirror or video image to enable cooperating subjects to position their own eyes within the field of view of a single narrow-angle camera.

Image focus assessment is performed in real-time (faster than video frame rate) by measuring spectral power in middle and upper frequency bands of the two-dimensional (2D) Fourier spectrum of each image frame and seeking to maximize this quantity either by moving an active lens or by providing audio feedback to subjects to adjust their range appropriately. The video rate execution speed of focus assessment (i.e., within 15 msec) is achieved by using a bandpass 2D filter kernel requiring only summation and differencing of pixels, and no multiplications, within the 2D convolution necessary to estimate power in the selected 2D spectral bands. Details are provided in Section 9.

Images passing a minimum focus criterion are then analyzed to find the iris, with precise localization of its boundaries using a coarse-to-fine strategy terminating in single-pixel precision estimates of the center coordinates and radius of both the iris and the pupil. Although the results of the iris
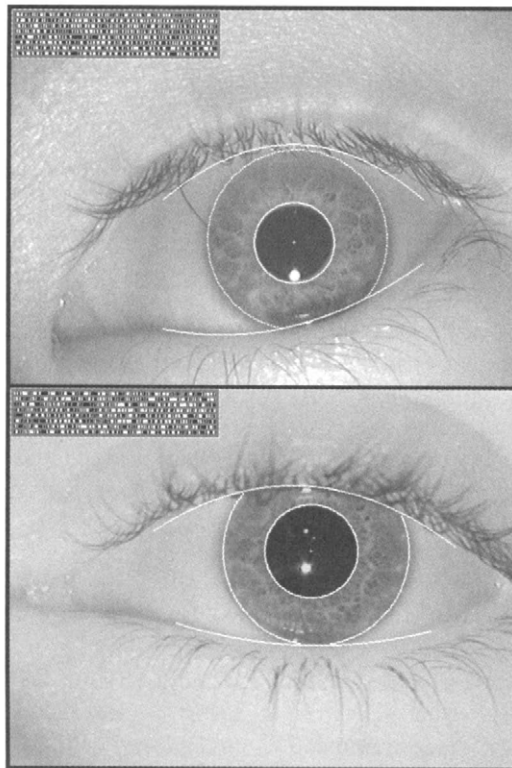

FIGURE 1 Examples of human iris patterns, imaged monochromatically at a distance of about 35 cm. The outline overlays show the results of the iris and pupil localization and eyelid detection steps. The bit streams pictured are the result of demodulation with complex-valued two-dimensional Gabor wavelets (2) to encode the phase sequence of each iris pattern.

search greatly constrain the pupil search, concentricity of these boundaries cannot be assumed. Very often the pupil center is nasal, and inferior, to the iris center. Its radius can range from 0.1 to 0.8 of the iris radius. Thus, all three parameters defining the pupillary circle must be estimated separately from those of the iris. A very effective integrodifferential operator for determining these parameters is

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (1)$$

where $I(x, y)$ is an image such as Fig. 1 containing an eye. The operator searches over the image domain (x, y) for the maximum in the blurred partial derivative with respect to increasing radius r, of the normalized contour integral of $I(x, y)$ along a circular arc ds of radius r and center coordinates (x_0, y_0). The symbol $*$ denotes convolution and $G_\sigma(r)$ is a smoothing function such as a Gaussian of scale σ. The complete operator behaves as a circular edge detector, blurred at a scale set by σ, searching iteratively for the maximal contour integral derivative at successively finer scales of analysis through the three parameter space of center coordinates and radius (x_0, y_0, r) defining a path of contour integration.

The operator in (1) serves to find the pupillary boundary and the outer (limbus) boundary of the iris, although the initial search for the limbus also incorporates evidence of an interior pupil to improve its robustness since the limbic boundary itself usually has extremely soft contrast when long-wavelength NIR illumination is used. Once the coarse-to-fine iterative searches for both these boundaries have reached single-pixel precision, then a similar approach to detecting curvilinear edges is used to localize both the upper and lower eyelid boundaries. The path of contour integration in (1) is changed from circular to arcuate, with spline parameters fitted by statistical estimation methods to model each eyelid boundary. Images with less than 50% of the iris visible between the fitted eyelid splines are deemed inadequate (e.g., in blink). The result of all these localization operations is the isolation of iris tissue from other image regions, as illustrated in Fig. 1 by the graphic overlay on the eye.

Because pupils are generally not exactly round, the inner boundary of the iris coordinate system should not be forced to be a circle. Instead, once the approximate circular boundary has been determined, an active contour ("snake") is allowed to deform into the true pupillary boundary by seeking equilibrium between internal (smoothness) and external (edge data) "energy" terms. A general principle in model-based computer vision is that where the data are strong, weak constraints should be used, whereas strong constraints should be used where the data are weak. For irises imaged in infrared illumination as illustrated in Fig. 1, the outer boundary of the iris with the sclera (the limbus) is often a very weak signal, sometimes having only about 1% contrast.

Such weak-edge data should be localized under the strong constraint that this contour will be a circle. But the pupillary boundary offers such strong edge data of high contrast that it can be determined reliably with weak constraints, and so an active snake with only weak smoothness constraints is allowed to fit its true contour for setting the inner boundary of the pseudo-polar coordinate system.

3 Iris Feature Encoding by Two-Dimensional Wavelet Demodulation

Each isolated iris pattern is then demodulated to extract its phase information using quadrature 2D Gabor wavelets [6, 7, 11]. This encoding process is illustrated in Fig. 2. It amounts to a patch-wise phase quantization of the iris pattern, by identifying in which quadrant of the complex plane each resultant phasor lies when a given area of the iris is projected onto complex-valued 2D Gabor wavelets:

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

where $h_{\{Re, Im\}}$ can be regarded as a complex-valued bit whose real and imaginary parts are either 1 or 0 (sgn) depending on the sign of the 2D integral; $I(\rho, \phi)$ is the raw iris image in a dimensionless polar coordinate system that is size- and translation-invariant, and also corrects for pupil dilation as
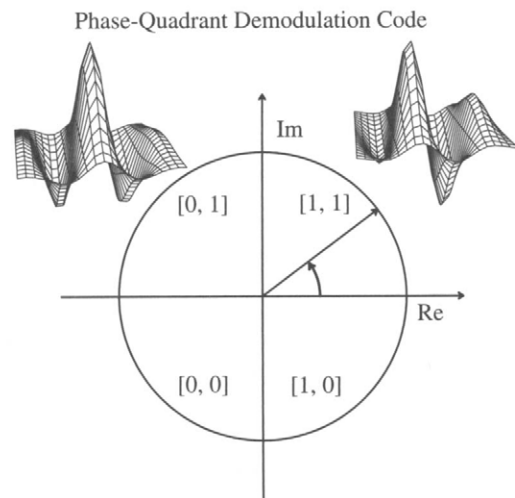

FIGURE 2 The phase demodulation process used to encode iris patterns. Local regions of an iris are projected (2) onto quadrature two-dimensional (2D) Gabor wavelets, generating complex-valued coefficients whose real and imaginary parts specify the coordinates of a phasor in the complex plane. The angle of each phasor is quantized to one of the four quadrants, setting two bits of phase information. This process is repeated all across the iris with many wavelet sizes, frequencies, and orientations, to extract 2,048 bits.

explained in a later section; α and β are the multiscale 2D wavelet size parameters, spanning an 8-fold range from 0.15 to 1.2 mm on the iris; ω is wavelet frequency, spanning three octaves in inverse proportion to β; and (r_0, θ_0) represent the polar coordinates of each region of iris for which the phasor coordinates $h_{\{Re, Im\}}$ are computed. Such phase quadrant coding sequences are illustrated for two irises by the bit streams shown graphically in Fig. 1. A desirable feature of the phase code definition given in Fig. 2 is that it is a cyclic, or gray code: In rotating between any adjacent phase quadrants, only a single bit changes, unlike a binary code in which two bits may change, making some errors arbitrarily more costly than others. Altogether 2,048 such phase bits (256 bytes) are computed for each iris, but in a major improvement over the author's earlier [8] algorithms, now 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 and thus should be ignored in the demodulation code as artifact.

The 2D Gabor wavelets were chosen for the extraction of iris information because of the nice optimality properties of these wavelets. Following the Heisenberg uncertainty principle as it applies generally to mathematic functions, filters that are well localized in frequency are poorly localized in space (or time), and vice versa. The 2D Gabor wavelets have the maximal joint resolution in the two domains simultaneously [6,7], which means that both "what" and "where" information about iris features is extracted with optimal simultaneous resolution. A further nice property of 2D Gabor wavelets is that because they are complex-valued, they allow the definition and assignment of phase variables to any point in the image.

Only phase information is used for recognizing irises because amplitude information is not very discriminating, and it depends upon extraneous factors such as imaging contrast, illumination, and camera gain. The phase bit settings which code the sequence of projection quadrants as shown in Fig. 2 capture the information of wavelet zero-crossings, as is clear from the sign operator in (2). The extraction of phase has the further advantage that phase angles remain defined regardless of how poor the image contrast may be, as illustrated by the extremely out-of-focus image in Fig. 3. Its phase bit stream has statistical properties such as run lengths similar to those of the codes for the properly focused eye images in Fig. 1. (Fig. 3 also illustrates the robustness of the iris- and pupil-finding operators, and the eyelid detection operators, despite poor focus.) The benefit which arises from the fact that phase bits are set also for a poorly focused image as shown here, even if based only on random CCD thermal noise, is that different poorly focused irises never become confused with each other when their phase codes are compared. By contrast, images of different faces look increasingly alike when poorly resolved, and can be confused with each other by appearance-based face recognition algorithms.
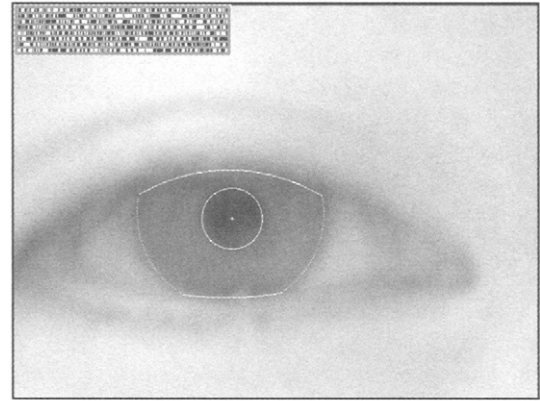

FIGURE 3 Illustration that even for poorly focused eye images, the bits of a demodulation phase sequence are still set, primarily by random CCD noise. This prevents poorly focused eye images from being falsely matched, as they may be in amplitude-based representations.

4 The Test of Statistical Independence: Combinatorics of Phase Sequences

The key to iris recognition is the failure of a test of statistical independence, which involves so many degrees of freedom that this test is virtually guaranteed to be passed whenever the phase codes for two different eyes are compared, but to be uniquely failed when any eye's phase code is compared with another version of itself.

The test of statistical independence is implemented by the simple Boolean Exclusive-OR operator (XOR) applied to the 2,048 bit phase vectors that encode any two iris patterns, masked (AND'ed) by both of their corresponding mask bit vectors to prevent noniris artifacts from influencing iris comparisons. The XOR operator $\otimes$ detects disagreement between any corresponding pair of bits, while the AND operator $\cap$ ensures that the compared bits are both deemed to have been uncorrupted by eyelashes, eyelids, specular reflections, or other noise. The norms ($\| \, \|$) of the resultant bit vector and of the AND'ed mask vectors are then measured in order to compute a fractional hamming distance (HD) as the measure of the dissimilarity between any two irises, whose two phase code bit vectors are denoted {*codeA*, *codeB*} and whose mask bit vectors are denoted {*maskA*, *maskB*}:

$$\text{HD} = \frac{\|(codeA \otimes codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|} \quad (3)$$

The denominator tallies the total number of phase bits that mattered in iris comparisons after artifacts such as eyelashes and specular reflections were discounted, so the resulting HD is a fractional measure of dissimilarity; 0 would represent a perfect match. The Boolean operators $\otimes$ and $\cap$ are applied in vector form to binary strings of length up to the word length of the Computer Processing Unit (CPU), as a single machine

instruction. Thus for example on an ordinary 32-bit machine, any two integers between 0 and 4 billion can be XOR'ed in a single machine instruction to generate a third such integer, each of whose bits in a binary expansion is the XOR of the corresponding pair of bits of the original two integers. This implementation of (3) in parallel 32-bit chunks enables extremely rapid comparisons of iris codes when searching through a large database to find a match. On a 300-MHz CPU, such exhaustive searches are performed at a rate of about 100,000 irises per second; on a 3-GHz server, about a million iris comparisons can be performed per second.

Because any given bit in the phase code for an iris is equally likely to be 1 or 0, and different irises are uncorrelated, the expected proportion of agreeing bits between the codes for two different irises is HD = 0.500. The histogram in Fig. 4 shows the distribution of HDs obtained from 9.1 million comparisons between different pairings of iris images acquired by licensees of these algorithms in the United Kingdom, the United States, Japan, and Korea. There were 4,258 different iris images, including ten each of one subset of 70 eyes. Excluding those duplicates of (700×9) same-eye comparisons, and not double-counting pairs, and not comparing any image with itself, the total number of unique pairings between different eye images whose HDs could be computed was $[(4,258 \times 4,257 - 700 \times 9)/2] = 9,060,003$. Their observed mean HD was $p = 0.499$ with standard deviation $\sigma = 0.0317$; their full distribution in Fig. 4 corresponds to a binomial having $N = p(1 - p)/\sigma^2 = 249$ df, as shown by the solid curve. The extremely close fit of the theoretical binomial to the observed distribution is a consequence of the fact that each comparison between two phase code bits from two

different irises is essentially a Bernoulli trial, albeit with correlations between successive "coin tosses."

In the phase code for any given iris, only small subsets of bits are mutually independent due to the internal correlations, especially radial, within an iris. (If all $N = 2048$ phase bits were independent, then the distribution in Fig. 4 would be very much sharper, with an expected standard deviation of only $\sqrt{p(1 - p)/N} = 0.011$ and so the HD interval between 0.49 and 0.51 would contain most of the distribution.) Bernoulli trials that are correlated [18] remain binomially distributed but with a reduction in N, the effective number of tosses, and hence an increase in the σ of the normalized HD distribution. The form and width of the HD distribution in Fig. 4 tell us that the amount of difference between the phase codes for different irises is distributed equivalently to runs of 249 tosses of a fair coin (Bernoulli trials with $p = 0.5$, $N = 249$). Expressing this variation as a discrimination entropy [5] and using typical iris and pupil diameters of 11 and 5 mm, respectively, the observed amount of statistical variability among different iris patterns corresponds to an information density of about 3.2 bits/mm² on the iris.

The theoretical binomial distribution plotted as the solid curve in Fig. 4 has the fractional functional form

$$f(x) = \frac{N!}{m!(N - m)!} p^m (1 - p)^{(N-m)} \qquad (4)$$

where $N = 249$, $p = 0.5$, and $x = m/N$ is the outcome fraction of N Bernoulli trials (e.g., coin tosses that are "heads" in each run). In our case, x is the HD, the fraction of phase bits that happen to agree when two different irises are compared. To validate such a statistical model we must also study the behaviour of the tails, by examining quantile-quantile plots of the observed cumulatives versus the theoretically predicted cumulatives from 0 up to sequential points in the tail. Such a "Q-Q" plot is given in Fig. 5. The straight line relationship reveals very precise agreement between model and data, over a range of more than three orders of magnitude. It is clear from Figs. 4 and 5 that it is extremely improbable that two different irises might disagree by chance in fewer than at least a third of their bits. (Of the 9.1 million iris comparisons plotted in the histogram of Fig. 4, the smallest HD observed was 0.334.) Computing the cumulative of $f(x)$ from 0 to 0.333 indicates that the probability of such an event is about 1 in 16 million. The cumulative from 0 to just 0.300 is 1 in 10 billion. Thus, even the observation of a relatively poor degree of match between the phase codes for two different iris images (say, 70% agreement or HD = 0.300) would still provide extraordinarily compelling evidence of identity, because the test of statistical independence is still failed so convincingly.

I also compared genetically identical eyes in the same manner, in order to discover the degree to which their textural patterns were correlated and hence genetically determined.
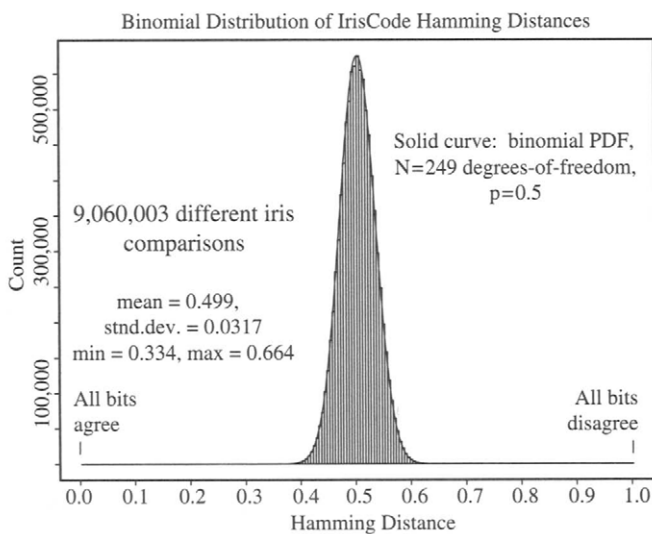

FIGURE 4 Distribution of hamming distances from all 9.1 million possible comparisons between different pairs of irises in the database. The histogram forms a perfect binomial distribution with $p = 0.5$ and $N = 249$ degrees-of-freedom, as shown by the solid curve (4). The data implies that it is extremely improbable for two different irises to disagree in less than about a third of their phase information.

Figure text:

Binomial Distribution of IrisCode Hamming Distances

9,060,003 different iris comparisons

Solid curve: binomial PDF, N=249 degrees-of-freedom, p=0.5

mean = 0.499,
stnd.dev. = 0.0317
min = 0.334, max = 0.664

All bits agree

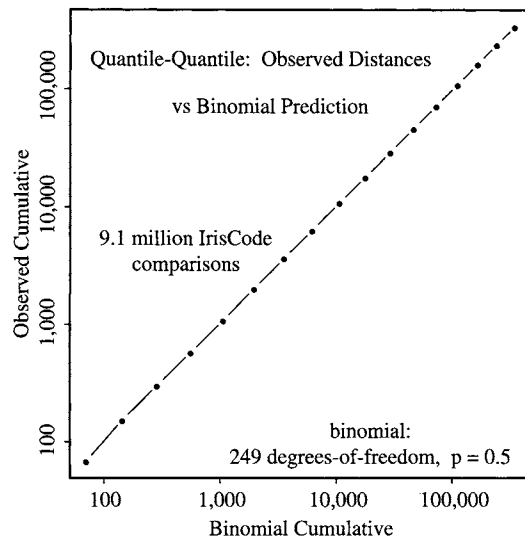All bits disagree

Count

Hamming Distance

FIGURE 5 Quantile-quantile plot of the observed cumulatives under the left tail of the histogram in Fig. 4 versus the predicted binomial cumulatives. The close agreement over several orders of magnitude strongly confirms the binomial model for phase bit comparisons between different irises.
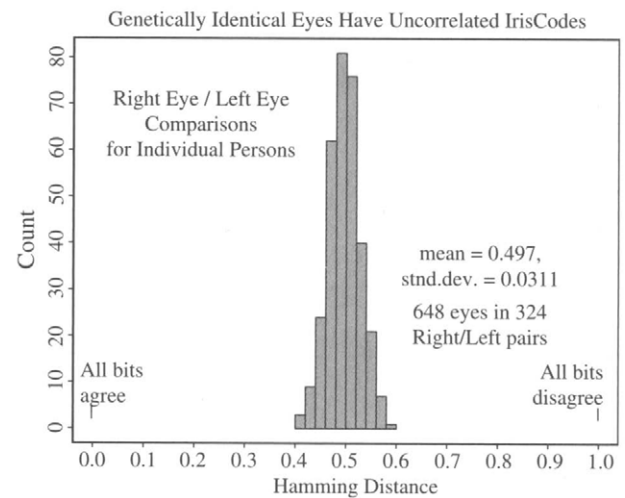

FIGURE 6 Distribution of hamming distances between genetically identical irises, in 648 paired eyes from 324 persons. The data are statistically indistinguishable from that shown in Fig. 4 comparing unrelated irises. Unlike eye color, the phase structure of iris patterns therefore appears to be epigenetic, arising from random events and circumstances in the morphogenesis of this tissue.

A convenient source of genetically identical irises are the right and left pair from any given person; such pairs have the same genetic relationship as the four irises of monozygotic twins, or indeed the prospective $2N$ irises of N clones. Although eye colour is of course strongly determined genetically, as is overall iris appearance, the detailed patterns of genetically identical irises appear to be as uncorrelated as they are among unrelated eyes. Using the same methods as described above, 648 right/left iris pairs from 324 persons were compared pairwise. Their mean HD was 0.497 with a standard deviation of 0.031, and their distribution (Fig. 6) was statistically indistinguishable from the distribution for unrelated eyes (Fig. 4). A set of six pairwise comparisons among the eyes of actual monozygotic twins also yielded a result (mean $HD = 0.507$) expected for unrelated eyes. It appears that the phenotypic random patterns visible in the human iris are almost entirely epigenetic [12].

5 Recognizing Irises Regardless of Size, Position, and Orientation

Robust representations for pattern recognition must be invariant to changes in the size, position, and orientation of the patterns. In the case of iris recognition, this means we must create a representation that is invariant to the optical size of the iris in the image (which depends upon the distance to the eye, and the camera optical magnification factor); the size of the pupil within the iris (which introduces a non-affine pattern deformation); the location of the iris within the image; and the iris orientation, which depends upon head tilt,

torsional eye rotation within its socket (cyclovergence), and camera angles, compounded with imaging through pan/tilt eye-finding mirrors that introduce additional image rotation factors as a function of eye position, camera position, and mirror angles. Fortunately, invariance to all of these factors can readily be achieved.

For on-axis but possibly rotated iris images, it is natural to use a projected pseudo polar coordinate system. The polar coordinate grid is not necessarily concentric, since in most eyes the pupil is not central in the iris; it is not unusual for its nasal displacement to be as much as 15%. This coordinate system can be described as doubly dimensionless: The polar variable, angle, is inherently dimensionless, but in this case the radial variable is also dimensionless, because it ranges from the pupillary boundary to the limbus always as a unit interval [0, 1]. The dilation and constriction of the elastic meshwork of the iris when the pupil changes size is intrinsically modelled by this coordinate system as the stretching of a homogeneous rubber sheet, having the topology of an annulus anchored along its outer perimeter, with tension controlled by an (off-centered) interior ring of variable radius.

The homogeneous rubber-sheet model assigns to each point on the iris, regardless of its size and pupillary dilation, a pair of real coordinates (r, θ) where r is on the unit interval [0, 1] and θ is angle [0, 2π]. This normalization, or remapping of the iris image $I(x, y)$ from raw cartesian coordinates (x, y) to the dimensionless nonconcentric polar coordinate system (r, θ) can be represented as

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \qquad (5)$$

where $x(r, \theta)$ and $y(r, \theta)$ are defined as linear combinations of both the set of pupillary boundary points $(x_p(\theta), y_p(\theta))$ and the set of limbus boundary points along the outer perimeter of the iris $(x_s(\theta), y_s(\theta))$ bordering the sclera, both of which are detected by finding the maximum of the operator (1).

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_s(\theta) \qquad (6)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_s(\theta) \qquad (7)$$

Since the radial coordinate ranges from the iris inner boundary to its outer boundary as a unit interval, it inherently corrects for the elastic pattern deformation in the iris when the pupil changes in size.

The localization of the iris and the coordinate system described above achieve invariance to the 2D position and size of the iris, and to the dilation of the pupil within the iris. However, it would not be invariant to the orientation of the iris within the image plane. The most efficient way to achieve iris recognition with orientation invariance is not to rotate the image itself using the Euler matrix, but rather to compute the iris phase code in a single canonical orientation and then to compare this very compact representation at many discrete orientations by cyclic scrolling of its angular variable. The statistical consequences of seeking the best match after numerous relative rotations of two iris codes are straightforward. Let $f_0(x)$ be the raw density distribution obtained for the HDs between different irises after comparing them only in a single relative orientation; for example, $f_0(x)$ might be the binomial defined in (4). 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 HD acceptance criterion x:

$$F_0(x) = \int_0^x f_0(x)dx \qquad (8)$$

or, equivalently,

$$f_0(x) = \frac{d}{dx}F_0(x) \qquad (9)$$

Clearly, then, the probability of not making a false match when using criterion x is $1 - F_0(x)$ after a single test, and it is $[1 - F_0(x)]^n$ after carrying out n such tests independently at n different relative orientations. It follows that the probability of a false match after a "best of n" test of agreement, when using HD criterion x, regardless of the actual form of the raw

unrotated distribution $f_0(x)$, is:

$$F_n(x) = 1 - [1 - F_0(x)]^n \qquad (10)$$

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

$$f_n(x) = \frac{d}{dx}F_n(x)$$
$$= nf_0(x)[1 - F_0(x)]^{n-1} \qquad (11)$$

Each of the 9.1 million pairings of different iris images whose HD distribution was shown in Fig. 4, was submitted to further comparisons in each of seven relative orientations. This generated 63 million HD outcomes, but in each group of seven associated with any one pair of irises, only the best match (smallest HD) was retained. The histogram of these new 9.1 million best HDs is shown in Fig. 7. Since only the smallest value in each group of seven samples was retained, the new distribution is skewed and biased to a lower mean value (HD = 0.458), as expected from the theory of extreme value sampling. The solid curve in Fig. 7 is a plot of (11), incorporating (4) and (8) as its terms, and it shows an excellent fit between theory (binomial extreme value sampling) and data. The fact that the minimum HD observed in all of these millions of rotated comparisons was about 0.33 illustrates the extreme improbability that the phase sequences for two different irises might disagree in fewer than a third of their bits. This suggests that to identify people by their iris
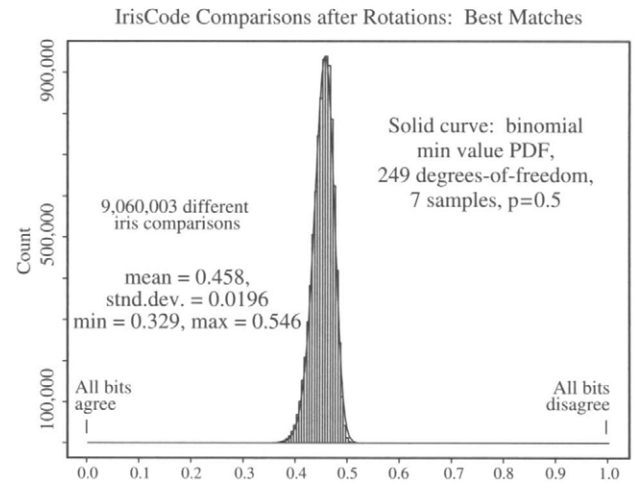

FIGURE 7   Distribution of hamming distances for the same set of 9.1 million comparisons shown in Fig. 4, but allowing for seven relative rotations and preserving only the best match found for each pair. This "best of n" test skews the distribution to the left and reduces its mean from about 0.5 to 0.458. The solid curve is the theoretical prediction for such "extreme-value" sampling, as described by (4) and (8) through (11).

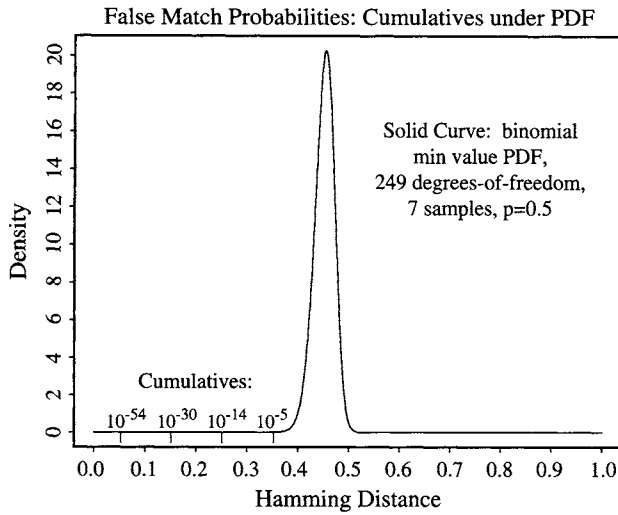False Match Probabilities: Cumulatives under PDF

FIGURE 8 Calculated cumulatives under the left tail of the distribution seen in Fig. 7, up to sequential points, using the functional analysis described by (4) and (8) – (11). The extremely rapid attenuation of these cumulatives reflects the binomial combinatorics that dominate (4). This accounts for the astronomic confidence levels against a false match, when executing this test of statistical independence.

TABLE 1 Cumulatives under (11) giving single false match probabilities for various HD criteria

HD Criterion	Odds of False Match
0.26	1 in 10^{13}
0.27	1 in 10^{12}
0.28	1 in 10^{11}
0.29	1 in 13 billion
0.30	1 in 1.5 billion
0.31	1 in 185 million
0.32	1 in 26 million
0.33	1 in 4 million
0.34	1 in 690,000
0.35	1 in 133,000

patterns with high confidence, we need to demand only a very forgiving degree of match (say, HD ≤ 0.32).

6 Uniqueness of Failing the Test of Statistical Independence

The statistical data and theory presented in the previous section show that we can perform iris recognition successfully just by a test of statistical independence. Any two different irises are statistically "guaranteed" to pass this test of independence; and any two images that fail this test must be images of the same iris. Thus, it is the unique failure of the test of independence, that is the basis for iris recognition.

It is informative to calculate the significance of any observed HD matching score, in terms of the likelihood that it could have arisen by chance from two different irises. These probabilities give a confidence level associated with any recognition decision. Figure 8 shows the false match probabilities marked off in cumulatives along the tail of the distribution presented in Fig. 7 (same theoretical curve (11) as plotted in Fig. 7 and with the justification presented in Figs. 4 and 5.) Table 1 enumerates false match probabilities, the cumulatives of (11), as a more fine-grained function of HD decision criterion between 0.26 and 0.35.

Calculation of the large factorial terms in (4) was done with Stirling's approximation, which errs by less than 1% for $n \ge 9$:

$$n! \approx \exp(n\ln(n) - n + \frac{1}{2}\ln(2\pi n)) \qquad (12)$$

The practical importance of the astronomic odds against a false match when the match quality is better than about HD ≤ 0.32, as shown in Fig. 8 and in Table 1, is that such high confidence levels allow very large databases to be searched exhaustively without succumbing to any of the many opportunities for suffering a false match. The requirements of operating in one-to-many "identification" mode are vastly more demanding than operating merely in one-to-one "verification" mode (in which an identity 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 clearly P_N, the probability of making at least one false match when searching a database of N unrelated patterns, is

$$P_N = 1 - (1 - P_1)^N \qquad (13)$$

because $(1 - P_1)$ is the probability of not making a false match in single comparisons; this must happen N independent times; and so $(1 - P_1)^N$ is the probability that such a false match never occurs.

It is interesting to consider how a seemingly impressive biometric one-to-one "verifier" would perform in exhaustive search mode once databases become larger than about 100, in view of (13). For example, a face recognition algorithm that truly achieved 99.9% correct rejection when tested on non-identical faces, hence making only 0.1% false matches, would seem to be performing at a very impressive level because it must confuse no more than 10% of all identical twin pairs (since about 1% of all persons in the general population have an identical twin). But even with its $P_1 = 0.001$, how good would it be for searching large databases?

Using (13) we see that when the search database size has reached merely $N = 200$ unrelated faces, the probability of at least one false match among them is already 18%. When the search database is just $N = 2000$ unrelated faces, the

probability of at least one false match has reached 86%. Clearly, identification is vastly more demanding than one-to-one verification, and even for moderate database sizes, merely "good" verifiers are of no use as identifiers. Observing the approximation that $P_N \approx N P_1$ for small $P_1 \ll \frac{1}{N} \ll 1$, 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.

The algorithms for iris recognition exploit the extremely rapid attenuation of the HD distribution tail created by binomial combinatorics, to accommodate very large database searches without suffering false matches. The HD threshold is adaptive, to maintain $P_N < 10^{-6}$ regardless of how large the search database size N is. As Table 1 illustrates, this means that if the search database contains 1 million different iris patterns, it is only necessary for the HD match criterion to adjust downwards from 0.33 to 0.27 in order to maintain still a net false match probability of 10^{-6} for the entire database.

7 Decision Environment for Iris Recognition

The overall "decidability" of the task of recognizing persons by their iris patterns is revealed by comparing the HD distributions for same versus for different irises. The left distribution in Fig. 9 shows the HDs computed between 7,070 different pairs of same-eye images at different times, under different conditions, and usually with different cameras; and the right distribution gives the same 9.1 million comparisons among different eyes shown earlier. To the degree that one can confidently decide whether an observed sample belongs to the left or the right distribution in Fig. 9, iris recognition can be

FIGURE 10 The decision environment for iris recognition under very favorable conditions, using the same camera, distance, and lighting.

successfully performed. Such a dual distribution representation of the decision problem may be called the "decision environment," because it reveals the extent to which the two cases (same versus different) are separable and thus how reliably decisions can be made, since the overlap between the two distributions determines the error rates.

Whereas Fig. 9 shows the decision environment under less favourable conditions (images acquired by different camera platforms), Fig. 10 shows the decision environment under ideal (almost artificial) conditions. Subjects' eyes were imaged in a laboratory setting using the same camera with fixed zoom factor and at fixed distance, and with fixed illumination. Not surprisingly, more than half of such image comparisons achieved an HD of 0.00, and the average HD was a mere 0.019. It is clear from comparing Figs. 9 and 10 that the "authentics" distribution for iris recognition (the similarity between different images of the same eye, as shown in the left-side distributions), depends very strongly upon the image acquisition conditions. However, the measured similarity for "imposters" (the right-side distribution) is almost completely independent of imaging factors. Instead, it just reflects the combinatorics of Bernoulli trials, as bits from independent binary sources (the phase codes for different irises) are compared.

For two-choice decision tasks (e.g., same versus different), such as biometric decision making, the "decidability" index d' is one measure of how well separated the two distributions are, since recognition errors would be caused by their overlap. If their two means are μ_1 and μ_2, and their two standard deviations are σ_1 and σ_2, then d' is defined as

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}} \qquad (14)$$

This measure of decidability is independent of how liberal or conservative is the acceptance threshold used. Rather, by
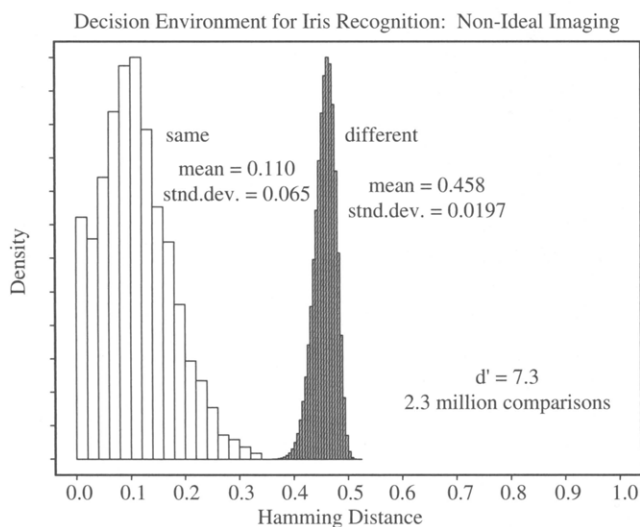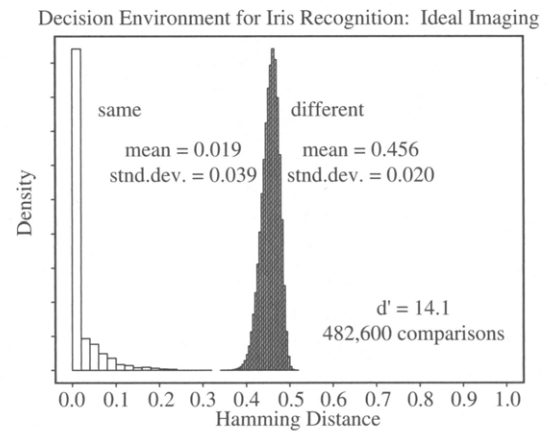

FIGURE 9 The decision environment for iris recognition under relatively unfavorable conditions, using images acquired at different distances, and by different optical platforms.

measuring separation, it reflects the degree to which any improvement in (say) the false match error rate must be paid for by a worsening of the failure-to-match error rate. The performance of any biometric technology can be calibrated by its d' score, among other metrics. The measured decidability for iris recognition is $d' = 7.3$ for the nonideal (crossed platform) conditions presented in Fig. 9, and it is $d' = 14.1$ for the ideal imaging conditions presented in Fig. 10.

Based on the left-side distributions in Figs. 9 and 10, one could calculate a table of probabilities of failure to match, as a function of HD match criterion, just as in Table 1 for false match probabilities based on the right-side distribution. However, such estimates may not be stable because the "authentics" distributions depend strongly on the quality of imaging (e.g., motion blur, focus, noise, etc.) and would be different for different optical platforms. As illustrated earlier by the badly defocused image of Fig. 3, phase bits are still set randomly with binomial statistics in poor imaging, and so the right distribution is the stable asymptotic form both in the case of well imaged irises (Fig. 10) and poorly imaged irises (Fig. 9). Imaging quality determines how much the same-iris distribution evolves and migrates leftward, away from the asymptotic different-iris distribution on the right. In any case, we note that for the 7,070 same-iris comparisons shown in Fig. 9, their highest HD was 0.327, which is below the smallest HD of 0.329 for the 9.1 million comparisons between different irises. Thus, a decision criterion slightly below 0.33 for the empiric data sets shown can perfectly separate the dual distributions. At this criterion, using the cumulatives of (11) as tabulated in Table 1, the theoretical false match probability is 1 in 4 million.

Notwithstanding this diversity among iris patterns and their apparent singularity because of so many dimensions of random variation, their utility as a basis for automatic personal identification would depend upon their relative stability over time. There is a popular belief that the iris changes systematically with one's health or personality, and even that its detailed features reveal the states of individual organs ("iridology"); but such claims have been discredited (e.g., [3, 17]) as medical fraud. In any case, the recognition principle described here is intrinsically tolerant of a large proportion of the iris information being corrupted, say up to about a third, without significantly impairing the inference of personal identity by the simple test of statistical independence.

8 Speed Performance Summary

On a low-cost 300-MHz reduced instruction set (RISC) processor, the execution times for the critical steps in iris recognition are as shown in Table 2, using optimized integer code:

The search engine can perform about 100,000 full comparisons between different irises per second on each such a CPU, or 1 million in about a second on a 3-GHz server, because of

TABLE 2 Execution speeds of various stages in the iris recognition process on a 300-MHz RISC processor

Operation	Time
Assess image focus	15 msec
Scrub specular reflections	56 msec
Localize eye and iris	90 msec
Fit pupillary boundary	12 msec
Detect and fit both eyelids	93 msec
Remove lashes and contact lens edges	78 msec
Demodulation and IrisCode creation	102 msec
XOR comparison of two IrisCodes	10 msec

the efficient implementation of the matching process in terms of elementary Boolean operators $\otimes$ and $\cap$ acting in parallel on the computed phase bit sequences. If a database contained many millions of enrolled persons, then the inherent parallelism of the search process should be exploited for the sake of speed by dividing up the full database into smaller chunks to be searched in parallel. The confidence levels shown in Table 1 indicate how the decision threshold should be adapted for each of these parallel search engines to ensure that no false matches were made despite several large-scale searches being conducted independently. The mathematics of the iris recognition algorithms, particularly the binomial-class distributions (4) (11) that they generate when comparing different irises, make it clear that databases the size of an entire country's population could be searched in parallel to make confident and rapid identification decisions using parallel banks of inexpensive CPUs, if such iris code databases existed.

9 Appendix: Two-Dimensional Focus Assessment at the Video Frame Rate

The acquisition of iris images in good focus is made difficult by the optical magnification requirements, the restrictions on illumination, and the target motion, distance, and size. All of these factors act to limit the possible depth of field of the optics, because they create a requirement for a lower F number to accommodate both the shorter integration time (to reduce motion blur) and the light dilution associated with long focal length. The iris is a 1-cm target within a roughly 3-cm–wide field that one would like to acquire at a range of about 30 cm to 50 cm, and with a resolution of about five line pairs per mm. In a fixed-focus optical system, the acquisition of iris images almost always begins in poor focus. It is therefore desirable to compute focus scores for image frames very rapidly, either to control a moving lens element or to provide audible feedback to the subject for range adjustment, or to select which of several frames in a video sequence is in best focus.

Optical defocus can be fully described as a phenomenon of the 2D Fourier domain. An image represented as a 2D function of the real plane, $I(x, y)$, has a 2D Fourier transform $F(\mu, v)$ defined as:

$$F(\mu, v) = \frac{1}{(2\pi)^2} \int \int I(x, y) \exp(-i(\mu x + v y)) dx dy \quad (15)$$

In the image domain, defocus is normally represented as convolution of a perfectly focused image by the 2D point-spread function of the defocused optics. This point-spread function is often modelled as a Gaussian whose space constant is proportional to the degree of defocus. Thus, for perfectly focused optics, this optical point-spread function shrinks almost to a delta function, and convolution with a delta function has no effect on the image. Progressively defocused optics equates to convolving with ever wider point-spread functions.

If the convolving optical point-spread function causing defocus is an isotropic Gaussian whose width represents the degree of defocus, it is clear that defocus is equivalent to multiplying the 2D Fourier transform of a perfectly focused image with the 2D Fourier transform of the "defocusing" (convolving) Gaussian. This latter quantity is itself just another 2D Gaussian within the Fourier domain, and its spread constant there (σ) is the reciprocal of that of the image-domain convolving Gaussian that represented the optical point-spread function. Thus the 2D Fourier transform $D_\sigma(\mu, v)$ of an image defocused by degree $1/\sigma$ can be related to $F(\mu, v)$, the 2D Fourier transform of the corresponding perfectly focused image, by a simple model such as:

$$D_\sigma(\mu, v) = \exp\left(-\frac{\mu^2 + v^2}{\sigma^2}\right) F(\mu, v) \quad (16)$$

This expression reveals that the effect of defocus is to attenuate primarily the highest frequencies in the image and that lower frequency components are affected correspondingly less, since the exponential term approaches unity as the frequencies (μ, v) become small. (For simplicity, this analysis has assumed isotropic optics and isotropic blur, and the optical point-spread function has been described as a Gaussian just for illustration. But the analysis can readily be generalized to non-Gaussian and to anisotropic optical point-spread functions.)

This spectral analysis of defocus suggests that an effective way to estimate the quality of focus of a broadband image is simply to measure its total power in the 2D Fourier domain at higher spatial frequencies, since these are the most attenuated by defocus. One may also perform a kind of "contrast normalization" to make such a spectrally-based focus measure independent of image content, by comparing the ratio of

power in higher frequency bands to that in slightly lower frequency bands. Such spectrally based measurements are facilitated by exploiting Parseval's theorem for conserved total power in the two domains:

$$\int \int |I(x, y)|^2 dx dy = \int \int |F(\mu, v)|^2 d\mu dv \quad (17)$$

Thus, high-pass filtering an image, or bandpass filtering it within a ring of high spatial frequency (requiring only a 2D convolution in the image domain), and integrating the power contained in it, is equivalent to computing the actual 2D Fourier transform of the image (a more costly operation) and performing the corresponding explicit measurement in the selected frequency band. Since the computational complexity of a fast Fourier transform on $n \times n$ data is $\mathcal{O}(n^2 \log_2 n)$, some 3 million floating-point operations are avoided which would be otherwise be needed to compute the spectral measurements explicitly. Instead, only about 6,000 integer multiplications per image are needed by this algorithm, and no floating-point operations. Computation of focus scores is based only on simple algebraic combinations of pixel values within local closed neighborhoods, repeated across the image.

Pixels are combined according to the following (8×8) convolution kernel:

-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	+3	+3	+3	+3	-1	-1
-1	-1	+3	+3	+3	+3	-1	-1
-1	-1	+3	+3	+3	+3	-1	-1
-1	-1	+3	+3	+3	+3	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1

The simple weights mean that the sum of the central (4×4) pixels can just be tripled, and then the outer 48 pixels subtracted from this quantity; the result is squared and accumulated as per (17); and then the kernel moves to the next position in the image, selecting every fourth row and fourth column. This highly efficient discrete convolution has a simple 2D Fourier analysis.

The kernel shown is equivalent to the superposition of two centered square box functions, one of size (8×8) and amplitude -1, and the other one of size (4×4) and amplitude $+4$. (For the central region in which they overlap, the two therefore sum to $+3$.) The 2D Fourier transform of each of these square functions is a 2D "sinc" function, whose size parameters differ by a factor of two in each of the dimensions and whose amplitudes are equal but opposite, since the two component boxes have equal but opposite volumes. Thus the overall kernel has a 2D Fourier transform $K(\mu, v)$
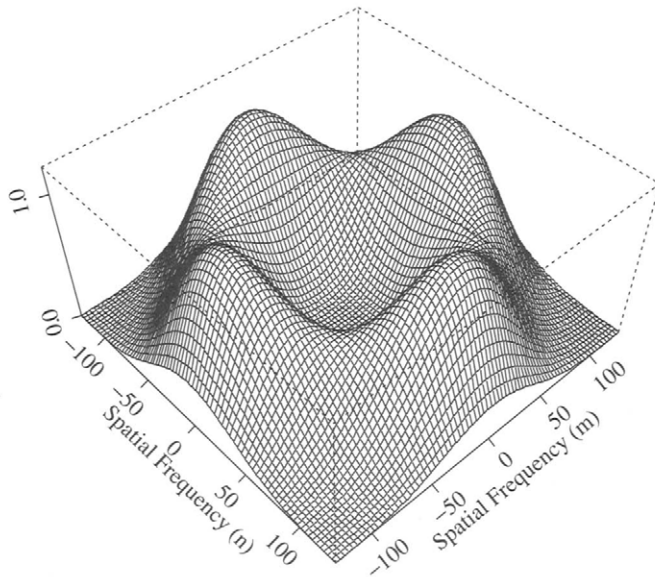
FIGURE 11 The two-dimensional Fourier power spectrum of the convolution kernel used for rapid focus assessment.

which is the difference of two, differently-sized, 2D sinc functions:

$$K(\mu, \nu) = \frac{\sin(\mu)\sin(\nu)}{\pi^2 \mu \nu} - \frac{\sin(2\mu)\sin(2\nu)}{4\pi^2 \mu \nu} \qquad (18)$$

The square of this function of μ and ν in the 2D Fourier domain is plotted in Fig. 11, revealing $K^2(\mu, \nu)$, the convolution kernel's 2D power spectrum.

Clearly, low spatial frequencies (near the center of the power spectral plot in Fig. 11) are ignored, reflecting the fact that the pixel weights in the convolution kernel all sum to zero, while a bandpass ring of upper frequencies are selected by this filter. The total power in that band is the spectral measurement of focus. Finally, this summated 2D spectral power is passed through a compressive nonlinearity of the form: $f(x) = 100 \cdot x^2/(x^2 + c^2)$ (where parameter c is the half-power corresponding to a focus score of 50%), to generate a normalized focus score in the range of 0 to 100 for any image. The complete execution time of this 2D focus assessment algorithm, implemented in C using pointer arithmetic, operating on a (480×640) image, is 15 msec on a 300-MHz RISC processor.

References

[1] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 721–732 (1997).

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class-specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 711–720 (1997).

[3] L. Berggren, "Iridology: A critical review," *Acta Ophthalmologica* 63, 1–8 (1985).

[4] M. R. Chedekel, "Photophysics and photochemistry of melanin," in *Melanin: Its Role in Human Photoprotection* (Valdenmar, Overland Park, 1995), 11–23.

[5] T. Cover and J. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[6] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A* 2, 1160–1169 (1985).

[7] J. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust. Speech Signal Process.* 36, 1169–1179 (1988).

[8] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.* 15, 1148–1161 (1993).

[9] J. Daugman, U. S. Patent No. 5,291,560: *Biometric Personal Identification System Based on Iris Analysis.* US Government Printing Office, Washington DC, (1994).

[10] J. Daugman, "Statistical richness of visual phase information: Update on recognizing persons by their iris patterns," *Intl. J. Comput. Vis.* 45, 25–38 (2001).

[11] J. Daugman and C. Downing, "Demodulation, predictive coding, and spatial vision," *J. Opt. Soc. Amer. A* 12, 641–660 (1995).

[12] J. Daugman and C. Downing, "Epigenetic randomness, complexity, and singularity of human iris patterns," *Proc. Royal Society: Biological Sciences* 268, 1737–1740 (2001).

[13] P. Kronfeld, "Gross anatomy and embryology of the eye," in H. Davson, ed. *The Eye.* (Academic Press, London, 1962).

[14] A. Pentland and T. Choudhury, "Face recognition for smart environments," *Computer* 33, 50–55 (2000).

[15] P. J. Phillips, A. Martin, C. L. Wilson, and M. Przybocki, "An introduction to evaluating biometric systems," *Computer* 33, 56–63 (2000).

[16] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 1090–1104 (2000).

[17] A. Simon, D. M. Worthen, and J. A. Mitas, "An evaluation of iridology," *J. Amer. Med. Assoc.* 242, 1385–1387 (1979).

[18] R. Viveros, K. Balasubramanian, and N. Balakrishnan, "Binomial and negative binomial analogues under correlated Bernoulli trials," *Am. Statistician* 48, 243–247 (1984).