Iris image segmentation and sub-optimal images

James R. Matey\textsuperscript{a},*, Randy Broussard\textsuperscript{a}, Lauren Kennell\textsuperscript{b}

\textsuperscript{a}Center for Biometric Signal Processing, ECE Department, Maury Hall, MS 14B, US Naval Academy, Annapolis, MD 21402-5025, USA
\textsuperscript{b}Johns Hopkins Applied Physics Laboratory, MS 24-W211, 11100 Johns Hopkins Road, Laurel, MD 20723-6099, USA

\begin{abstract}
Iris recognition is well developed and works well for optimal or near-optimal iris images. Dealing with sub-optimal images remains a challenge. Resolution, wavelength, occlusion and gaze are among the most important factors for sub-optimal images. In this paper, we explore the sensitivity of matching to these factors through analysis and numerical simulation, with particular emphasis on the segmentation portion of the processing chain.
\end{abstract}

\section{Background}

The iris is the colored portion of the eye that resides between the white sclera and the nominally black pupil. The iris is rich in detail and the detail comes about from processes that have sensitive dependence on initial conditions, so that these details, though stable for a given eye, are random. The reader's two eyes, directed at this page, have identical genetics; they will likely have the same color and may well show some large scale pattern similarities; nevertheless, they have quite different iris pattern details. Use of the detailed patterns of the iris as an identifier was likely first suggested by Bertillon \cite{1,2} in the late 1800s. However, it was not until the mid-1990s that advances in computer vision, image sensors and computers enabled a practical implementation of this idea by Daugman \cite{3}.

Iris recognition is now one of the most accurate and effective means of biometric identification. The United Arab Emirates Expellees Tracking and Border Control System \cite{4} is an outstanding example of the technology. As of late 2008, IrisGuard reported \cite{5} the results in Table 1.

Since Daugman's original paper, there have been many suggestions for alternative iris recognition algorithms. Bowyer et al. \cite{6} recently presented an excellent review of these methods. However, at this time, essentially all of the large scale implementations of iris recognition are based on the Daugman iris recognition algorithms \cite{7}; the most widely used implementation is usually referred to as iris2pi. iris2pi accepts an image of the form seen in Fig. 1 and generates a template using a process similar\textsuperscript{1} to the following:

\begin{itemize}
  \item Find the pupil/iris and iris/sclera boundaries.
  \item Extract the iris from the image.
  \item Remap the iris into doubly dimensionless pseudo-polar coordinates, with the horizontal axis representing angle from 0 to $2\pi$, and the vertical axis representing radial distance from the pupil boundary to the scleral boundary.
  \item Establish an $8 \times 128$ grid on the remapped image; 8 locations along the radial direction and 128 along the angular.
  \item At each of the grid points measure the local phase in a predetermined bandwidth by executing a dot product between the remapped image and a pair of sine-like and cosine-like Gabor wavelets, forming a ratio of the two and taking the arc-tangent of the result.
  \item Digitize the resulting angle to two bits, corresponding to the quadrant in which the local phase resides.
  \item Assemble the bits into an array 8 bits high and 256 bits wide. This is the phase part of an iris2pi template.
  \item As the phase bits are computed, estimate the quality of the bits. In a mask array that shadows the phase array, set the bit to 1 if the corresponding phase bit can be trusted, set the bit to zero if it cannot.
  \item The resulting template has 256 bytes of phase information and 256 bytes of mask information.
\end{itemize}

\textsuperscript{1} This prescription is one possible implementation of the published algorithm; this prescription is chosen for pedagogical reasons. The internal details of the commercial iris2pi algorithms are proprietary and almost certainly differ from this prescription. Code based on this prescription is unlikely to be as fast as code based on optimized prescriptions used in commercial implementations.
The process is illustrated in Figs. 1–3. The determination of the iris/sclera and iris/pupil boundaries is frequently referred to as segmentation – the topic of this special issue; it is the first step in the process and, as we shall see, arguably the most important – and most difficult.

To compare two iris2pi templates, A and B, Daugman uses a fractional Hamming distance, frequently shortened to Hamming distance and abbreviated HD. The comparison process is

- For each phase bit in A, pick up the corresponding phase bit in B and the corresponding mask bits in A and B.
- If either mask bit is not set, do nothing – go onto the next phase bit in A.
- If both mask bits are set, increment a bits compared counter and compare the phase bits.
- If the phase bits from A and B are the same, do nothing and move onto the next phase bit in A.
- If the phase bits differ, increment a bits different counter and then move onto the next phase bit in A.
- When all the bits in A are exhausted, form the ratio of the bits different and the bits compared counters. This ratio is the fractional Hamming distance.

Daugman has presented compelling evidence [8] that the binomial distribution, with \( \sim 250 \) degrees of freedom, is a good model for the distribution of the phase bits in an iris2pi template. Using that model, the probability that fewer than \( \sim 0.32 \) of the bits disagree for two independent iris2pi templates is of the order of \( 1 \times 10^{-6} \); that fewer than \( \sim 0.25 \) disagree has a probability of the order of \( 1 \times 10^{-12} \).

It is important to recognize that the remapped image and the resulting templates represent cylindrical surfaces rather than flat surfaces – the left and right edges represent 0 and \( 2\pi \). This is crucial when we consider what happens if the image for template B is rotated slightly about the pupil center relative to that for template A. Circular rotation of an iris image about the pupil center (assuming circular pupils and irises and concentric irises and pupils) is equivalent to a barrel shift (a rotation of a cylinder on its axis) of its template. Fig. 4 shows the relationship between fractional Hamming distance and barrel shift for the template of Fig. 1, compared to itself. Note that a shift of one angular position \( \pi \) increases the HD from zero to approximately 0.25. Shifts of 5 positions or more produce HDs of approximately 0.5, equivalent to comparisons between unrelated iris images. The overshoot for shifts between 2 and 5 is likely the result of correlations resulting from

---

Table 1
Performance of the UAE expellee system as of November 2008.

<table>
<thead>
<tr>
<th>Database Size (templates)</th>
<th>&gt;1.6 million</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persons searched against database</td>
<td>&gt;20 million</td>
</tr>
<tr>
<td>Total cross comparisons</td>
<td>&gt;20 trillion</td>
</tr>
<tr>
<td>False matches</td>
<td>Zero</td>
</tr>
<tr>
<td>Persons caught</td>
<td>&gt;300 thousand</td>
</tr>
</tbody>
</table>

---

2 There are about 8 times that number of phase bits. The phase bits display correlation, particularly along the radial direction, as can be seen in Fig. 3.

3 The overhead per bit is \( \pi/128 \sim 0.05 \) rad, \( 360/128 \sim 2.8 \).
the interaction of the Gabor wavelets and the iris image structures on this scale.

The image that we used for this discussion, Fig. 1, shows an eye with no occlusion due to eyelid or eyelash. This enabled us to postpone consideration of a vital component of a commercially viable prescription – detection of eyelids and eyelashes. Eyelids, eyelashes and specularities will all obscure the iris. The obscured regions must be dealt with somehow in any algorithm. In iris2pi, obscured areas are flagged in the mask bits of the template. Specularities are generally easy to identify. Eyelids and eyelashes are much more difficult. Despite its importance, the literature on this topic comprises a small subset of the published work on iris recognition; Xu et al. [9], Kong and Zhang [10], Huang et al. [11], He et al. [12] and Bachoo and Tapamo [13] are among the more prominent members of the subset. A thorough discussion of this topic is beyond the scope of this paper.

2. Effects of segmentation errors

We are now in a position to consider the effects of segmentation errors. There are many types of segmentation errors; errors include:

- Pupil center error.
- Pupil radius error (or equivalent for non-circular pupil models).
- Iris center error.
- Iris radius error (or equivalent for non-circular iris models).
- Deviation of pupil or iris from the boundary model (e.g. deviation from circularity).
  - Missshapen iris.
  - Off-axis imaging.
- Errors related to occlusion.
  - Eyelid.
  - Eyelashes.
  - Specularities.

Let us consider a very simple error – a simple horizontal displacement of the pupil center from its ideal location – with the assumption that the iris and pupil are concentric to start. We will compute the average angular displacement of each location in the template and combine that with the data of Fig. 4 to provide an estimate of impact of such errors on the Hamming distance. The average angular displacement (radians) of points in the normalized image and the corresponding template can be modeled as

\[
\frac{1}{\pi (r_i^2 - r_p^2)} \int_0^{\pi/2} \int_{r_p}^{r_i} (\delta \sin(\theta) r \, dr \, d\theta) = \frac{4 \delta}{\pi r_i + r_p}
\]

where \( r_i \) is the iris radius, \( r_p \) is the pupil radius and \( \delta \) is the horizontal shift of the pupil and iris centers. We integrate the angular shift over the area of the iris and normalize by that area to get the average shift. We integrate over a single quadrant because we are interested in the average of the absolute value of the shift and all four quadrants give the same results with differing signs.

For the iris in Fig. 1, the iris radius is approximately 110 pixels and the pupil radius is approximately 40 pixels. The average angular displacement is therefore \( \sim 0.008 \) rad or \( \sim 0.5^\circ \) per pixel of horizontal iris-pupil displacement. From Fig. 4, the HD change per radian of angular displacement is \( \sim 0.25/(2\pi/128) \sim 5 \). Hence, we estimate that the HD change per pixel of displacement will be \( 5 \times 0.008 = 0.04 \) – a not insignificant change.

We can perform a numerical experiment to test this model. Fig. 5 illustrates the nature of the experiment. We clear the pupil in the original image to a gray level approximating the iris and then superimpose a new pupil, shifted horizontally with respect to the original, on the image. We can generate a new template from this image and compare it with the original template.

Fig. 6 illustrates the results for shifts from \( -20 \) to \( +20 \). The magnitude of the effect is approximately 0.02/pupil pixel shift – about half as large as that predicted by the simplified model. This is a

![Fig. 4](image-url) Fractional Hamming distance as a function of barrel shift for a template generated from Fig. 1.

![Fig. 5](image-url) Iris image of Fig. 1, with pupil cleared to a gray value approximating the iris and a new pupil superimposed with a shift of 15 pixels from the original. The segmentation of the modified image is shown by the white and black circles surrounding the pupil and iris. Approximately 5% of the pupil area is disturbed by this operation – iris pixels covered over by the new iris or pixels “made up” in the original pupil region.

![Fig. 6](image-url) Change in Hamming distance vs. pupil shift for images modified as in Fig. 5. The horizontal axis is the absolute value of the shift.
3. Survey of segmentation methods

The specifics of iris image acquisition are crucial to the design and performance of iris segmentation algorithms. Most current commercial iris acquisition systems acquire images from a small, known stand-off distance of the order of 10 cm with an on-axis (or nearly on-axis) presentation – the subject is essentially looking directly at the camera. Segmentation methods for such images can rely on an expected size (pixel diameter) and near-circularity of the pupil (inner) and limbic (outer) iris boundaries. Often the iris boundaries are conceptualized as nearly concentric circles (alternatively, the pupil mass can be thought of as a solid disk and the limbic boundary as a circle), so one often proceeds by searching for these shapes, for instance with circular edge detection. In some newer segmentation methods the process is taken a step further to correct for irregularly shaped boundaries.

Systems that allow larger stand-off distances and less constraint on the subject are under development; AOptix [15]; Honeywell [16], Hoyos [17], Retica [18] and Sarnoff [19,20] have all demonstrated systems that work beyond 1 m and Matey et al. prepared a review on such acquisition systems in a soon to be published book [21]. These systems are more subject-friendly than the traditional systems. However, the less constraint on the subject, the more robust the segmentation method must be. In these systems, the apparent size of the iris and pupil will show much more variability and the images will also be more subject to motion blur, inconsistent illumination, shadows, significant eyelid/eyelash obscuration, eyeglass effects (the frames and/or large glare areas from the lenses) and variation in subject gaze angles which makes the iris boundaries appear more elliptical than circular.

The iris boundaries have been modeled as circles, ellipses and more complicated shapes. Whatever shape model is used, the iris boundaries generally need to be represented as closed curves projecting behind occluding eyelids and eyelashes. This is necessary for any iris recognition algorithm that uses the pseudo-polar representation proposed by Daugman to normalize the iris prior to feature encoding. Though one can imagine the use of open curves, it makes sense to use closed curves because the iris has closed boundaries – closed curves are simply better physical models of the iris boundaries.

Publications on iris segmentation number in the hundreds. Table 2 provides a non-exhaustive list and taxonomy: an overview with illustrative examples of the most commonly employed ideas and techniques. Most published approaches have been developed and tested on image databases collected with traditional iris systems with cooperative subjects. The review by Bowyer et al. [6] is an excellent source of additional information and references.

4. The Daugman algorithm

Daugman’s recognition algorithm is used in all or nearly all current commercial iris recognition systems. Indeed, the integro-differential operator for circular edge detection, and the pseudo-polar coordinate transform, which are two of the image pre-processing steps introduced by Daugman in his first papers on this topic, have been incorporated into various other proposed recognition methods. Therefore it is natural to begin this section with the Daugman segmentation method [7].

To obtain a first approximation to the pupil boundary, limbic boundary, and eyelid boundary, the integro-differential operator

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

(2)

is applied, where $I(x, y)$ are the image grayscale values, $G_\sigma(r)$ is a smoothing function such as a Gaussian of scale $\sigma$, and the contour...
integral is along circles given by center $(x_0, y_0)$ and radius $r$. This operator finds the maximum blurred partial derivative of the image grayscale values with respect to a radial variable, of a contour integral along circles when searching for the pupil and limbic boundaries, and is modified to search along arcs for eyelid boundaries. In the most recent version of the algorithm [22], which accommodates off-axis images, the plots of gradients for the pupil and limbic boundaries form two “snakes,” and each of which is then approximated by a discrete Fourier series. The iris ring is thus bounded by smooth closed curves which project behind occluding eyelids, but in general neither boundary curve is exactly circular or elliptical. The last step in the segmentation process is the detection of pupil edges and regions overlapping the iris ring in unpredictable ways. Fortunately, the grayscale distribution in the upper half of the iris shows multimodal mixing; the eyelash pixels are eliminated by thresholding.

5. Other approaches

Many variations of, and alternatives to, Daugman’s segmentation method have been proposed, as illustrated in Table 2. We can extract a few common themes from these proposals:

- Combinations of binary morphology and thresholding to reduce noise and/or classify important regions within the image.

- Edge detection followed by thresholding or a Hough transform, where the edge detection operator should return an optimized or near-optimized value at the iris boundaries.

- Active contour methods for irregular iris boundaries, occluding artifacts and off-axis images.

Table 2

<table>
<thead>
<tr>
<th>Number</th>
<th>Reference</th>
<th>Modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Y. Du et al., A new approach to iris pattern recognition, in: SPIE European Symposium on Optics/Photonics in Defence and Security, London, UK, October 2004</td>
<td>BMT</td>
</tr>
<tr>
<td>20</td>
<td>A. Abhyankar, S. Schuckers, Active shape models for effective iris segmentation, Proceeding of SPIE 6202 (2006)</td>
<td>AC/ASM</td>
</tr>
</tbody>
</table>

BMT, binary morphology and thresholding; Stat, statistical patterns; EDT, edge detection and thresholding; IDO, integro differential operator; Hough, Hough transform; NNET, neural networks; Pol, polar coordinate search; FSD, Fourier spectral density; AC/ASM, active contour/active shape models.
that the limbic boundary is mapped exactly to a line only if the pupil and limbic centers are co-located.

There are a number of notable examples of edge detection methods. Wildes [25] proposed using a gradient based binarized image (an edge detected image) followed by a circular Hough transform to locate the center and radius of the maximum gradient; other variants of this method have been proposed [26–32]. Typically an image is blurred with a Gaussian function to reduce noise. Next, the image is edge detected using the Canny edge detector which is the method of choice. Last, the iris boundary is located using a circular or elliptical Hough transform. Since an elliptical Hough transform has a very large parameter search space, the search space is typically reduced by assuming the outer iris boundary has the same eccentricity as the pupil boundary. Du et al. proposed using a Sobel edge detector and looking for a straight line in the pseudo-polar image as an approximation to the circular Hough transform [33]; this approach is similar to that proposed by Camus and Wildes [34].

Other methodologies for edge detection based on statistical distributions and local image statistics [23] have also been proposed which use the fact that the pixels within the iris exhibit different grayscale distributions than those along the iris boundary, and therefore have different statistical properties.

Newer segmentation methods have employed more sophisticated edge detection and curve fitting, especially for locating the limbic boundary, eyelids, and other occlusions, since these problems are not well-resolved with edge detection operators being constrained by rigid circular or elliptical models. To that end, some of the updated methods use active contours for greater flexibility and robustness, for instance [24,35–37].

Alternatively, it has been proposed [38] to use artificial neural networks to classify each pixel as “iris” or “not iris” using local statistical moments, directional derivatives, and location of the pixel relative to a known pupil center as input features. The neural network classifier has the advantage of returning the iris mask directly, detecting and removing eyelids without curve fitting or shape models.

Segmentation is complicated. As noted earlier, it is likely the most difficult aspect of iris recognition. At the present state of our knowledge, there is not a single best segmentation method. The optimal segmentation method will vary depending on the details of the image and on the details of the template generation algorithm that follows the segmentation. As one example, the limbic boundary is soft and its softness depends on the wavelength of the light used to create the image. Hand segmentation of the limbic boundary can yield different results depending on the operator – and can yield different results with the same operator depending on the gamma adjustment of the display.

In light of these considerations, we may well ask, “What is a correct segmentation?” Gross errors are easily recognized; however, to our knowledge, there are no firm rules for distinguishing between two segmentations that are both plausible, but different.

One approach to testing of segmentation for non-ideal images was illustrated earlier in this paper. Given an ideal image that segments well, we can simulate non-ideal images by appropriate transforms, segment those images and then transform the resulting segmentation back to the space of the ideal image and compare the segmentation of the ideal image with the segmentations of the simulated non-ideal images.

6. Segmentation-free algorithms

It is possible to imagine iris recognition without segmentation – per se. There is an impetus to develop such algorithms because the first claim in the Daugman patent [3] is:

A method for uniquely identifying a particular human being by biometric analysis of the iris of the eye, comprising the following steps:

- acquiring an image of an eye of the human to be identified;
- isolating and defining the iris of the eye within the image, wherein said isolating and defining step includes the steps of:
  - defining a circular pupillary boundary between the iris and pupil portions of the image;
  - defining another circular boundary between the iris and sclera portions of the image, using arcs that are not necessarily concentric with the pupillary boundary;
  - establishing a polar coordinate system on the isolated iris image, the origin of the coordinate system being the center of the circular pupillary boundary, wherein the radial coordinate is measured as a percentage of the distance between the said circular pupillary boundary and said circular boundary between the iris and sclera;
  - defining a plurality of annular analysis bands within the iris image;
- analyzing the iris to generate a presenting iris code;
- comparing said presenting code with a previously generated reference iris code to generate a measure of similarity between said presenting iris code and said reference code;
- converting said similarity measure into a decision that said iris codes either do or do not arise from the same iris;
- calculating a confidence level for the decision.

All of the other claims in the patent derive from this claim. Hence, an algorithm that did not employ segmentation would likely not fall under this patent.

Since this paper is about segmentation, we will not cover segmentation free methods any further.

7. Examples – sub-optimal images

There are now several databases of sub-optimal images that can be used to test iris recognition algorithms for performance on images that do not conform to the nominal standards that we expect for high quality iris images. Proença and colleagues at the University of Beira Interior [39] have constructed the UBIRIS 1.0 and 2.0 iris databases; Jonathon Phillips from the NIST [40] and his collaborators have constructed the ICE and MBGC databases. All of these are available, with some restrictions, to the biometrics community.

Fig. 9 presents two images from UBIRIS 2.0 that illustrate three important issues associated with sub-optimal images:

- The images are taken in visible light, rather than near-IR.
- In the right image, the eye is significantly occluded by the eyelids.
- In the left image, the subject gaze is directed well to his right.

These images do not illustrate a fourth important issue – lack of resolution.

The details of any image depend on the wavelength of the light used to create it. As extreme cases, images created using X-rays or radio waves are certainly different from those created using visible light; astronomical images from Hubble, Chandra and the VLBA are examples. The distribution of melanin in the human eye gives the iris its color. Melanin is much more absorbing in the visible than the IR. Hence we expect that the details seen in a visible light iris image may differ significantly from those seen in a near IR image. Boyce [41,42] has presented data to support this view.
It is certainly possible to construct examples of patterns that are simply different when viewed with different wavelengths – and completely un-correlated. It remains to be seen to what extent the features in the iris in images taken at different wavelengths are correlated. In the absence of strong correlation, it would be impossible to match a blue light iris image to a near IR iris image.

Occlusion of the iris by eyelids, any other opaque medium, or strong specularities makes it impossible to recover information from the occluded region. Lack of information leads to difficulty in segmentation, and even if segmentation can be carried out, loss of information has an adverse impact on the quality of the match. Consider an iris which is 50% occluded, but for which segmentation and template generation has succeeded. For matches in which the number of bits compared differs significantly from a nominal 911, Daugman [8] has shown that the Hamming distance should be adjusted using the following equation:

\[
MHD = 0.5 - (0.5 - HD) \sqrt{N/911}
\]  

(3)

where HD is the fractional Hamming distance, MHD is the modified HD and N is the number of bits compared. Consider the effect of reducing N to N/2 for a match where the nominal N = 911 and HD = 0.30. The MHD changes from 0.30 to 0.36; the false match rate for a MHD of 0.30 is \( \approx 10^{-7} \ldots \) for 0.36 it is \( \approx 10^{-4} \).

The effect of off-axis gaze was demonstrated earlier using a simple model for gaze to the left or right. The left hand image in Fig. 9 is worse – the off-axis gaze is both to the right and up and the eyeball is rotated so that a significant portion of the iris is occluded.

8. Quo vadis?

In the previous section, we pointed out four major issues with sub-optimal images:

- Resolution
- Wavelength
- Occlusion
- Gaze

Current iris algorithms probe the structure of the iris at specific length scales. If an image does not have enough resolution to provide a non-aliased sampling of the structure, the algorithms will almost certainly fail. Our options are then to probe the structure at a coarser scale or to somehow improve the resolution of the image. Combining frames of video to achieve enhanced resolution has been used with good effect in other domains; it may prove useful here.

The question of how strongly correlated iris patterns are across wavelength is at present an open question. If there is no correlation, there is no hope. If there is correlation, standard statistical methods may prove fruitful.

Occlusion hides information. There seems to be no hope of getting that information, except from another image that is not occluded. The use of video techniques may again be fruitful.

Gaze is the issue that is most likely to yield to a direct attack at the level of single images. Given a suitable model for the eye, it may be possible to estimate gaze and correct for off-axis effects well beyond the \( \approx 25^\circ \) mark at which conventional algorithms begin to break down. However, there are limits; at sufficiently large angles, the eye begins to occlude itself; when head rotations are taken into account, other facial features can occlude the eye. For a head rotation of 90°, only one eye is visible and only half of the iris in that eye is visible.

Segmentation is a critical step in existing iris recognition algorithms, and processing methods can be applied in many, and always evolving, combinations to segment iris images. As images are acquired in increasingly diverse situations, the segmentation and recognition methods will have to keep pace accommodating them.

Acknowledgments

Our work would be difficult to impossible without the advice and support of our colleagues at the USNA and elsewhere. Our thanks to Imad Malhaus and Joe O’Carroll of IrisGuard for their statistics on the UAE expellees tracking and border control system; our thanks to LG-Iris for access to their iris recognition SDK and to Mohammed Murad, Tim Meyerhoff and Jun Hong of LG for their support of our use of the SDK. We also acknowledge numerous helpful discussions with John Daugman and Rob Ives’ continuing leadership of the Center for Biometrics Signal Processing.

Several colleagues offered constructive criticism of a preliminary draft. Special thanks to Kevin Bowyer, John Daugman, Rob Ives and Rick Wildes for catching errors and suggesting clarifications. Any remaining errors of commission or omission are the responsibility of the authors.

This work has been funded through the US Naval Academy Center for Biometric Signal Processing.

References