NIST Technical Note TN 932452

Statistics of Visual Features in the Human Iris

George W. Quinn
James R. Matey
Patrick Grother
Edward C. Watters III, M.D.

National Institute of
Standards and Technology
U.S. Department of Commerce

NIST Technical Note TN 932452

Statistics of Visual Features in the Human Iris

George W. Quinn, James R. Matey, Patrick Grother
gw@nist.gov, james.matey@nist.gov, patrick.grother@nist.gov
Information Technology Laboratory, Information Access Division, Image Group

Edward C. Watters III, M.D.
drecw3@gmail.com
Member of the Family LLC

This publication is available free of charge from:
www.nist.gov/publications/statistics-visual-features-human-iris

August 2021

U.S. Department of Commerce
Gina Raimondo, Secretary

National Institute of Standards and Technology
James K. Olthoff, performing the non-exclusive functions and duties of the Under Secretary of Commerce for Standards and Technology and Director, NIST

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology (NIST), nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

All data used in the research reported in this paper that was obtained from human subjects was collected under Institutional Review Board (IRB)-approved protocols. The analysis of that data was conducted under protocols reviewed and approved by the NIST Research Protection Office (RPO) or the NIST Institutional Review Board. The NIST protocols relevant for this paper are one of the following:

- ITL-16-0002 Data Collection in Support of Evaluation of Biometric Algorithms
- ITL-16-0026 Evaluations of Iris Recognition Technology: Iris Exchange: IREX-IX

Figures containing images of biometric features in this paper are either one of the following:

- Figures reproduced (with permission) from papers in the open literature.
- Figures containing images from subjects in the protocols above who provided permission for use of their images as part of the consent process.

Some of the research reported in this paper makes use of the Consolidated Multispectral Iris Dataset (CMID). The CMID was collected under the Consolidated Multispectral Iris Dataset Program, sponsored by the Central Intelligence Agency (CIA) and the Department of Defense, Technical Support Working Group (TSWG). The collections were carried out using IRB-approved protocols.

National Institute of Standards and Technology Technical Note TN 932452
Natl. Inst. Stand. Technol. Tech. Note TN 932452, 25 pages (August 2021)
CODEN: NTNOEF

This publication is available free of charge from:
www.nist.gov/publications/statistics-visual-features-human-iris

Abstract

John Daugman's popular IrisCode algorithm extracts feature information using complex-valued 2D Gabor wavelets, placing equal weight on all visible regions of the iris. Although research has demonstrated that the algorithm is highly accurate [1–6], the match decisions it makes are generally not well understood by the casual observer. Comparing irides using prominent local features (e.g., iris crypts, freckles, and contraction furrows) more closely aligns with the way humans naturally perform pattern recognition.

This paper presents statistical information on the distribution, distinctiveness, and permanence of visible features in the human iris. Such information can be used to develop a robust and accurate approach to matching local features in the iris. Two feature extraction algorithms are utilized to identify local features: 1) the scale-invariant feature transform (SIFT), and 2) The University Notre Dame's iris feature extractor. Preliminary results suggest matching based on local features holds promise, particularly when matching iris images acquired at visible wavelengths.

Comments and recommendations may be directed to gw@nist.gov.

Key words

iris recognition, human adjudication, forensics, visible features

Table of Contents

List of Figures

Glossary

Brushfield spots: White, gray, or pale yellow spots on the surface of the iris caused by an aggregation of connective tissue [7]. These spots are usually evenly spaced around the midperipheral or peripheral iris of many individuals with Down syndrome (though not exclusively) and are composed of the same tissue as Wolfflin nodules.

Collarette: The thickest region of the iris separating the pupillary (i.e., inner) portion from the ciliary (i.e., outer) portion [8].

Fuchs' Crypts: Openings in the anterior border layer of the iris and to a variable degree the iris stroma. Fuch's Crypts are normally close to the collarette but can also be found toward the periphery of the iris.

Exclusion: A decision in forensic science that a particular person is not the source of a biometric sample.

Freckle: Brown, yellow, or gray spots on the surface of the iris caused by an aggregation of melanin. There is no disturbance of the underlying stromal architecture and occur in 77 percent of the population. (See Nevi).

Nevi: Similar to freckles, these spots on the iris are secondary to a buildup of melanin pigment from clumps of melanocytes and can have a variety of colors. They tend to be larger than freckles and always cause distortion of the underlying stromal architecture. They occur in 5 percent of the population and in greater than 80 percent of the cases occur below the horizontal meridian of the iris.

Inclusion: A decision in forensic science that a particular person is the source of a biometric sample.

Inconclusive: A decision in forensic science indicating that it cannot be determined whether a particular person is the source of a biometric sample.

Iris: The colored, annular, part of the eye responsible for controlling the size of the pupil.

Iris2pi: A popular and widely utilized automated iris recognition algorithm developed by John Daugman.

Mated Comparison: A comparison between two samples of the same biometric characteristic.

Non-mated Comparison: A comparison between two samples that are not from the same biometric characteristic.

SIFT: The scale-invarient feature transform (SIFT) is a feature detection algorithm geared toward localizing, describing, and matching local features in images.

Local Feature: A prominent and distinctive characteristic limited to a particular region or neighborhood of an image.

Wölfflin nodules: Similar to Brushfield spots, aggregations of connective tissue on the periphery of the iris [9]. Unlike Brushfield spots, they tend to be fewer in number, located more peripherally and less distinct [10]. Often 10 - 20 are present in each eye.

1. Overview

Most state-of-the-art iris matchers generate comparison scores based on details of the iris texture that are not easily characterized or understood by the casual, untrained human observer. These algorithms also place equal weight on all visible portions of the iris. Yet the human eye is naturally attracted to prominent features and objects as they often convey the most useful information [11]. An iris recognition methodology based on such visible features (e.g., crypts, freckles) would more closely align with the way our human brains conduct pattern matching and may make explanations of matching to casual, untrained human observers easier.

This report performs some of the preliminary steps required to establish a matching approach using prominent local features. Some of the statistical properties of local features investigated include the following:

- the number of such features per eye,
- the radial distribution of such features,
- the circumferential distribution of such features, and
- the size distribution of such features.

This information can be used to address concerns raised by the National Academy of Sciences pertaining to the reliability and defensibility of biometric comparison decisions [12].

We make use of automated computer vision algorithms to extract features from a large set of iris images. The two feature extraction algorithms we explore are

- the scale invariant feature transformation (SIFT) keypoint detector described by Lowe et. al. [13], and
- the crypt and dark spot detector proposed and implemented by Chen et. al. [14] at Notre Dame University.

2. Rationale for Exploration of Visible Local Features

The following sections describe the rationale for the use of visible local features. In summary, the three primary drivers are

- the ability to explain matches to a non-technical audience,
- the ability to distinguish between *identification*, *exclusion* and *inconclusive* decisions [15], and

- the extraction of useful information from high-value images that cannot be processed by conventional iris matching algorithms.

2.1 Explanation of matches

As noted, conventional iris matching algorithms utilize details of the iris that are not easily interpretable by human examiners. Although studies have demonstrated the high accuracy of such algorithms [16–19], there is a need in the forensics community for matching protocols for the iris that are both explainable and have known error rates. Figure 1 demonstrates the way two fingerprints are typically matched, by pairing local minutia points. Figure 2 illustrates how a similar procedure might be used to match local features in two images of the same iris acquired six months apart. Although fingerprint examiner decisions are generally regarded as reliable, research into statistical models that can objectively quantify the distinctiveness of identification decisions is ongoing [20].

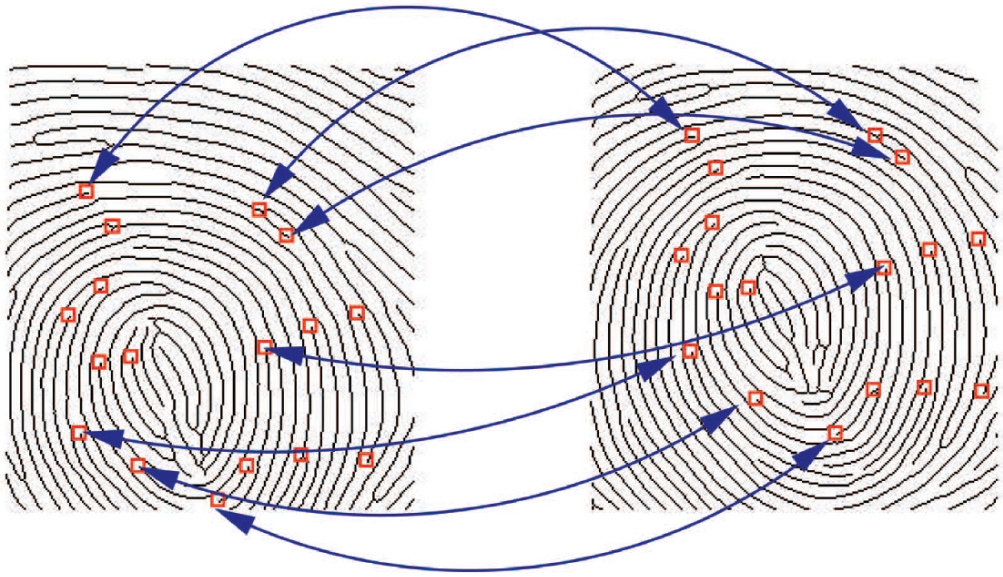


Fig. 1. Two images of a fingerprint illustrating corresponding minutiae (Figure 2 from Ross et. al.[21] with permission).

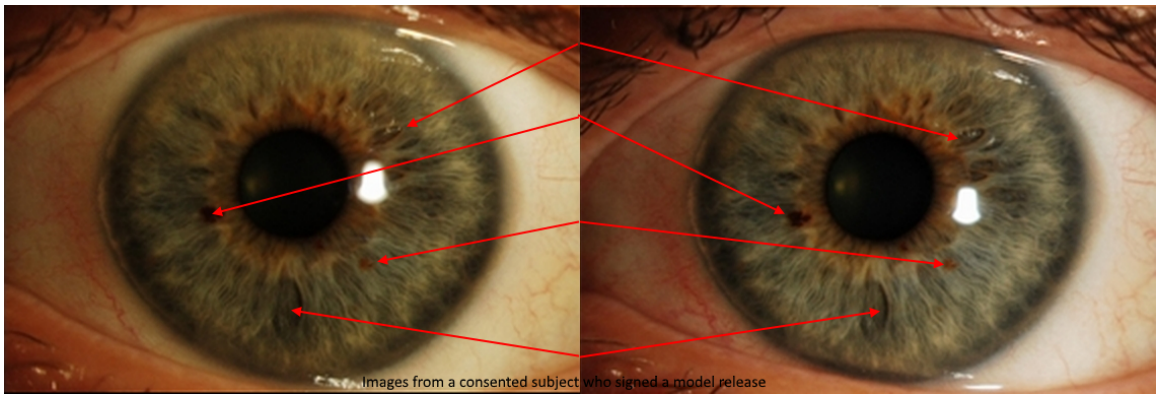


Fig. 2. Two images of an iris taken six months apart, illustrating corresponding visible local features. From top to bottom, the annotated features are 1) a crypt, 2) a freckle, 3) another freckle, and 4) a crypt.

2.2 Identifications, Exclusions, and Inconclusives

When a fingerprint examiner compares a fingerprint image from an unknown source (e.g., obtained from an object at a crime scene) with one from a known source (e.g., obtained from a suspect) they typically come to one of three conclusions [15]:

1. **Identification** - a positive identification is made (i.e., the examiner is confident that both fingerprints represent the same source).
2. **Exclusion** - a negative identification (i.e., the examiner is confident that the images represent different sources).
3. **Inconclusive** - the examiner cannot definitively make an identification or exclusion decision.

Fully automated methods of iris recognition are designed to make positive identifications and often do not distinguish between exclusions and inconclusives. Most deployed iris systems treat identification failures as inconclusives by default. For example, a user that was denied access on a first attempt usually has the opportunity to try again. Iris matchers produce measures of dissimilarity between iris samples. A low dissimilarity score is strong evidence that the images represent the same iris [17, 19, 22, 23]. DNA has also demonstrated that a single quantitative metric computed from population statistics can carry considerable weight in court, although iris recognition has not yet proved itself to be as robust as DNA. A high dissimilarity score, however, typically does not provide sufficient information to conclude whether an exclusion or inconclusive decision should follow. Operationally, a high dissimilarity score can occur if the quality of one or both of the iris images is low [18]. It can also occur if the images represent different irides.

Matching strategies based on pairing prominent local features can provide strong evidence for exclusion determinations. Most iris features are stable over time (e.g., Wölfflin nodules [9], Brushfield spots [7]) and do not form or disappear over a person's lifetime. For example, the presence of a freckle in one iris image but not in another when comparing the two images may conclude that the samples represent different persons.

That said, iris freckles can sometimes form or grow in prominence over a period of years and exposure to excessive amounts of light can accelerate the process [24]. Other health conditions (e.g., iritis, blunt trauma to the eye) can alter the appearance of the iris over a short period of time. Crypts can be distinctive but also tend to disappear when the pupil dilates. Circular contraction folds, on the other hand, often become more visible in when the pupil dilates. Existing knowledge in the fields of ophthalmology and biological anthropology should be leveraged when making a match decision, particularly when it comes to characterizing the prevalence, permanence, distinctiveness, and nature of specific local features.

Guidance on the proper acquisition of high quality iris images can be found in the IREX 5 documents [25].

3. Visible Feature Nomenclature

Table 1. A non-exhaustive compilation of visible iris feature definitions.

	Name	Description	Prevalence	Reference
1	Coloboma	Missing pieces of tissue in structures that form the iris.	1 in 10 000 people	[26]
2	Contraction Furrow	Fold in the iris tissue that forms a partial ring around the outer edge of the iris.	Common	[27]
3	Crypt	Typically diamond-shaped or oval-shaped holes in the iris tissue produced as the pupil opens and closes in response to light	Common	[28]
4	Ectopic Lentis et Pupillae	Displacement of both the pupil and lens from the normal anatomical position.	Uncommon	[29]
5	Freckle	A colored growth on or in the eye comprised of melanocytes.	77% of people	[30]
6	Hypoplasia	Any condition that causes an intact iris to erode or prevents an iris from developing properly.	Uncommon	[31]
7	Iridoplasty	A laser surgery procedure in which iris tissue in the far periphery is heated so that it will contract, causing a widening of the anterior chamber angle. This procedure is used to decrease the risk of angle closure glaucoma and occasionally is useful for treatment of angle closure glaucoma.	Uncommon	[32]
8	Iridectomy	The surgical removal of part of the iris.	Uncommon	[33]
9	Pupillary Ruff	The anterior continuation of the pigmented layer that lines the back surface of the iris to cover the edge of the pupil and slightly beyond. It is common to see a series of regular ridges in this area but this is subject to change as a result of aging and disease.	Common	[34]
10	Stroma	A delicate interlacement of fibrous tissue located immediately below the anterior border layer and comprises the bulk of the iris.	Common	[35]

4. Test Dataset

Testing in this report was performed using the Consolidated Multi-Spectral Iris Dataset (CMID) [36]. The dataset was collected under the CMID program, sponsored by the Central Intelligence Agency (CIA) and the Department of Defense Technical Support Working Group (TSWG). Copies of the dataset were provided to NIST for use in U.S. government iris recognition research under a data transfer agreement that restricts further distribution.

The dataset contains approximately 220 000 iris samples from more than 400 subjects collected at two geographic regions. The capture environment was highly controlled. During each capture session, images of the left and right irides were acquired across a range of wavelengths spanning from 400 nanometers to the infrared. Many subjects participated in multiple capture sessions on different days.

The actual test dataset used in this investigation is a subset of CMID consisting of 19 245 iris images from 233 subjects. These images were acquired at 800 nm (within the standard NIR range for iris cameras) and 620 nm (within the visible spectrum). Southern Methodist University (SMU) staff inspected the images to ensure serious problems were not present (e.g., the eye was not closed and there was adequate margin between the iris and the image edges). The images were paired to construct a set of 10 168 681 mated[1] and 57 17 334 nonmated[2] comparisons.

5. SIFT Algorithm

SIFT is an approach to detecting and describing local features in an image. SIFT was originally devised to recognize objects across different 3D views. It has since been applied to a range of computer vision problems, including automated biometric recognition [37–41]. Several studies have applied SIFT to iris recognition [42–47].

SIFT localizes stable points of interest referred to as *feature points* (a.k.a. keypoints). Several criteria are used to determine stability. Location stability is assessed using principal curvatures [48]. The ratio of the principal curvatures tends to be large along edges and lines, where feature points are poorly localized. For this reason, candidate points are disregarded when the ratio is greater than a preset value (10 in the original paper). Candidate points where the overall curvature is low are also disregarded as they tend to correspond to areas of low contrast, making them sensitive to noise

Figure 3 shows the locations of feature points found by SIFT in an iris using the param-

[1] A comparison is *mated* if the compared images represent the same iris

[2] A comparison is *nonmated* if the compared images represent different irides

eters listed in Appendix A. The image has been contrast-enhanced to improve the visibility of the texture. The feature points tend to be centered around dark spots (sometimes corresponding to crypts [7]), light spots (sometimes corresponding to Wölfflin nodules [9]), and areas where the stromal fibers branch and overlap.

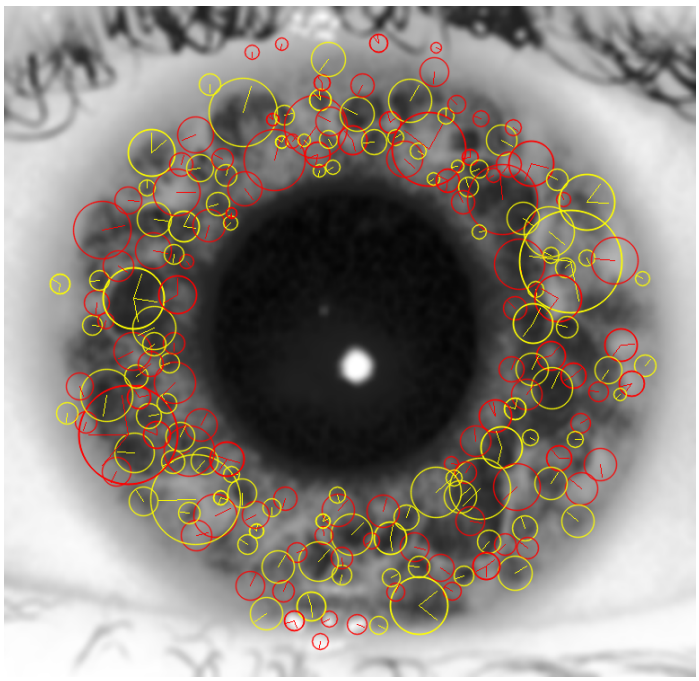


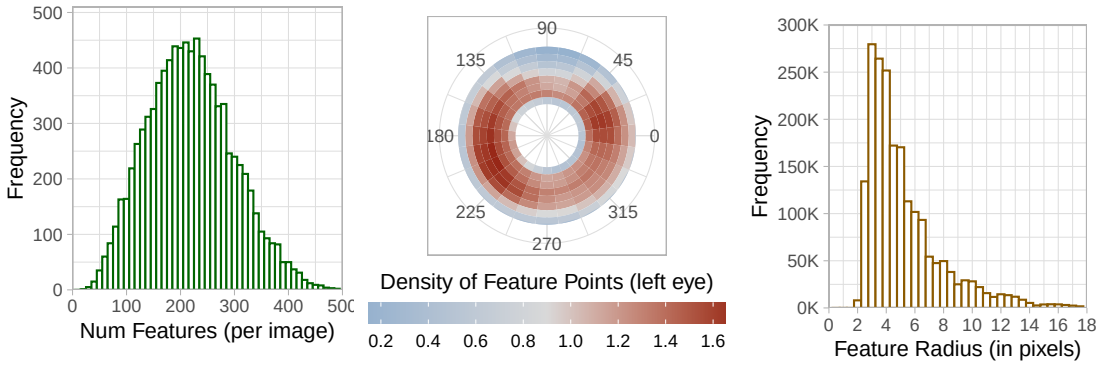
Fig. 3. Feature points located by SIFT in an iris image (contrast enhanced). The line segment indicate the gradient direction at the point. The color indicates the sign of the Laplacian at the feature point. Red will tend to correspond to light features while yellow will tend to correspond to dark features.

SIFT computes a descriptor for each feature point, represented as a 128-element feature vector. The values in the feature vector are computed using gradient orientations and magnitudes at locations in the immediate vicinity of the feature point. This gradient information is reduced into a 128-element feature vector in a way that ensures a degree of robustness to shape distortions and illumination changes. Lowe et al.'s approach is loosely based upon a model of biological vision [49] that is believed to mimic the way complex neurons in the primary visual cortex respond to gradients at particular orientations and spatial frequencies.

The dissimilarity between feature points is measured as the Euclidean distance between their descriptors. Lowe et al. state that the raw dissimilarity between feature points does not perform well because some descriptors are much more discriminative than others. They assert that a more effective metric is the ratio of the closest neighbor (among all of the feature points found in an image) and the second-closest neighbor. For object recognition, they recommend a threshold of 0.8, whereby a pairing is recognized as mated with high confidence

if the dissimilarity ratio is less than or equal to this threshold. SIFT only describes how to find the best pairing between images; it does not define an approach to quantifying the degree of similarity between images. That said, some papers have suggested that the number of successfully paired feature points between images is an effective metric [50, 51]. This approach closely parallels the way forensic fingerprint matching is performed in which an examiner attempts to pair minutia between samples and sometimes presents the number of paired minutia as a measure of the strength of a match.

5.1 Locations of SIFT Features



(a) Num. features found per iris image.

(b) Density plot of feature locations.

(c) Histogram of feature sizes.

Fig. 4. Basic statistical information on SIFT feature points.

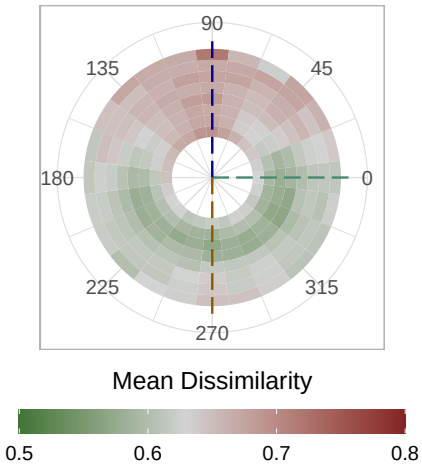
Figure 4 presents basic statistics on SIFT feature points found in iris images. SIFT finds an average of 241 ($\sigma = 72$) feature points per iris image with 90% of images containing between 122 and 360 feature points. That said, SIFT has a number of sensitivity parameters that can be adjusted to increase or decrease these numbers (see Appendix A).

Fewer feature points are found in the upper portion of the iris, largely because of upper eyelid occlusion. The lower quadrant of the iris (angle 225° to 315°) contains 60% more feature points than the upper quadrant (angle 45° to 135°) on average. Fewer features are also found in the outer periphery compared to the central region near the sphincter muscle.

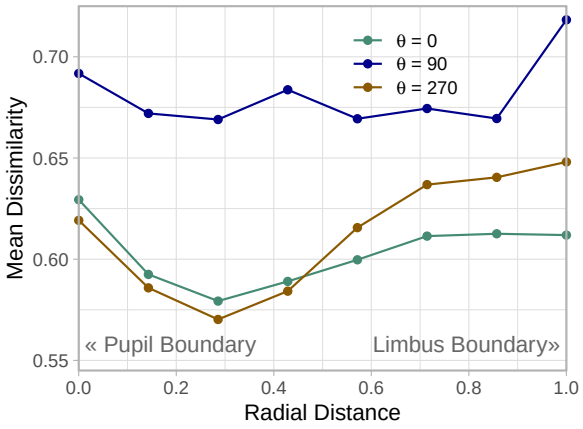
The distribution of feature point sizes is highly skewed toward smaller radii. The average feature point radius is 5.5 pixels, with 90 percent of feature points having radii between 2.6 and 12.1 pixels. Assuming a limbus radius of 10 mm and a common dilation ratio of 0.5, the average feature point radius would be 0.43 mm, with 90 percent of feature points having radii between 0.20 mm and 0.95 mm.

5.2 Stability of SIFT Features

5.2.1 Impact of Location



(a) Mean feature point dissimilarity by iris region.

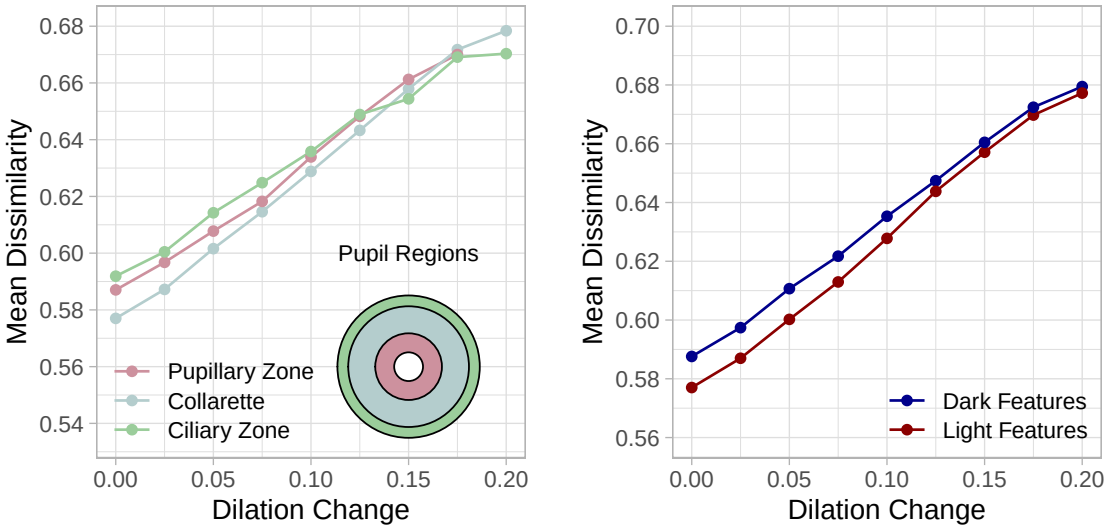


(b) Mean feature point dissimilarity as a function of distance from the pupil. Each curve represents a different polar angle.

Fig. 5. Stability of SIFT Features by location in the iris. The dissimilarity measure is the Euclidean distance between the feature vectors associated with each feature point. It is not a Hamming distance.

A stable feature should change little across different captures of the same iris, thereby producing low measures of dissimilarity when compared. The dissimilarity measure is the Euclidean distance between the two feature vectors, *not* a Hamming distance. Figure 5a plots mean dissimilarity for *mated* feature point comparisons in each region of the iris. Feature points in the upper portion of the iris appear to lack the stability of those in the lower portion. This is largely due to upper eyelid and eyelash interference. This may also explain why features in the inner portion of the iris seem to be more stable than those in the outer portion. Figure 5b shows that feature points tend to be the most stable roughly a quarter of the way out (corresponding to a radial distance of 0.25) from the pupil boundary. In summary, the most reliable feature points appear to be in the lower portion of the iris and closer to the pupil than the limbus.

5.2.2 Impact of Pupil Dilation



(a) Mean dissimilarity as a function of dilation change for feature points in different locations.

(b) Mean dissimilarity as a function of dilation change for ”light” and ”dark” feature points.

Fig. 6. Effect of dilation change on mean dissimilarity for different types of mated feature points. Dilation change was computed over the raw iris images; no image manipulation to simulate varying amounts of dilation was performed.

The SIFT algorithm was designed to be robust to shape distortions, although the developers focused on compensating for distortions caused by variations in viewing angle. Pupil dilation and constriction are difficult to compensate for, as crypts open and close, and fibrous tissue alters in orientation and shape. Figure 6a plots the average dissimilarity for mated feature points as a function of dilation change. The dilation change metric was first proposed in [17] and measures the degree to which the iris texture in one image must be radially stretched to match the iris in another image. The more one iris annulus has to be stretched to fit the thickness of the other iris annulus, the higher the dilation change value. Formally, dilation change is defined as

$$\Delta D = 1 - \left(\frac{R_i^2}{R_i^1}\right)\left(\frac{R_i^1 - R_p^1}{R_i^2 - R_p^2}\right) = 1 - \frac{1-D^1}{1-D^2} \tag{1}$$

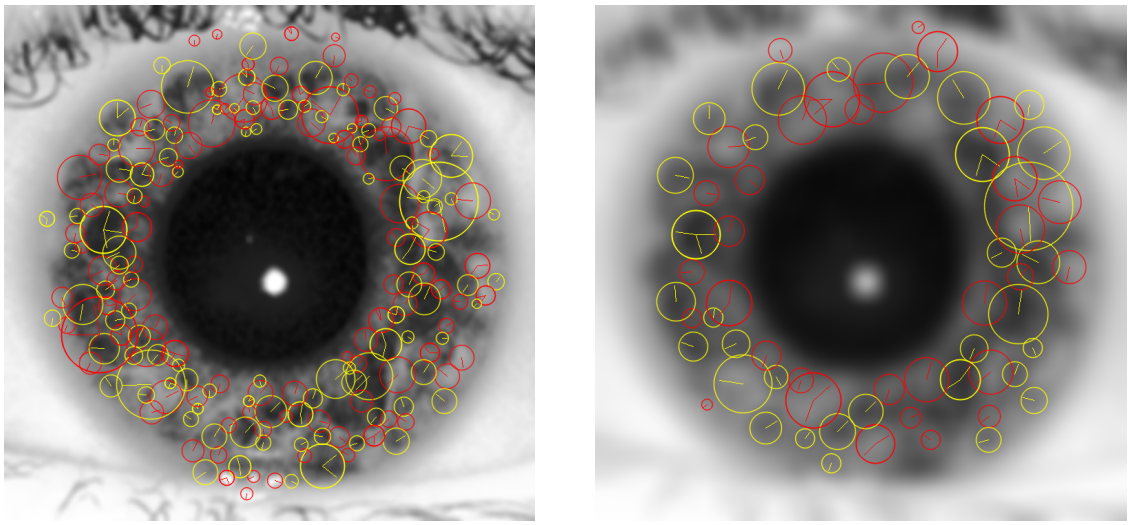
where R_p and R_i are estimates of the pupil and limbus radii respectively, $D = R_p/R_i$, and $D^1 > D^2$ is assumed without loss of generality. Dilation change is similar to, but not identical to, the difference in dilation ratios (i.e., $|D^1 - D^2|$) between the two iris images.

Figure 6a demonstrates that feature points in general have a tendency to be less recognizable as dilation change increases. Feature points in the vicinity of the collarette appear

slightly less recognizable than feature points in the pupillary and ciliary zone when the change in dilation is low. However, this does not appear to hold at higher amounts of dilation change. Figure 6b also demonstrates that dilation change detrimentally impacts the recognizability of feature points, but the effect is not appreciably different for ”light” features compared to ”dark” features. Generally speaking, the type and location of the feature point does not appear to have a significant impact on how it is affected by variations in pupil dilation.

5.2.3 Impact of Image Focus

Some of the images in the test dataset were acquired out-of-focus in which the finer details of the image are washed out and not visible. For these images, SIFT tends to only find the larger features of the iris. This is demonstrated in Figure 7, where the left image is in-focus and the right image is intentionally blurred to simulate defocus. Many more feature points are found in the in-focus image, 368 compared to just 79 in the blurred image. Additionally, only three of the feature points found in the blurred image have a radius smaller than three pixels compared to 93 in the in-focus image.



(a) In-focus **(b)** Out-of-focus

Fig. 7. An in-focus (left) and blurred (right) image with circles showing the locations of SIFT features. The color indicates whether the Laplacian at the feature is positive (yellow) or negative (red). Yellow tends to correspond to darker features while red tends to correspond to lighter features.

Figure 8 demonstrates that these observations hold true in general, with SIFT tending to find fewer feature points in out-of-focus images. Figure 9 further demonstrates that

differences in image sharpness reduce the similarity between *mated* feature points[3] feature points. Thus, not only are fewer feature point found in out-of-focus images, but those features that are found tend to be less distinctive. Image blur was measured using the standard S_3 metric [52].

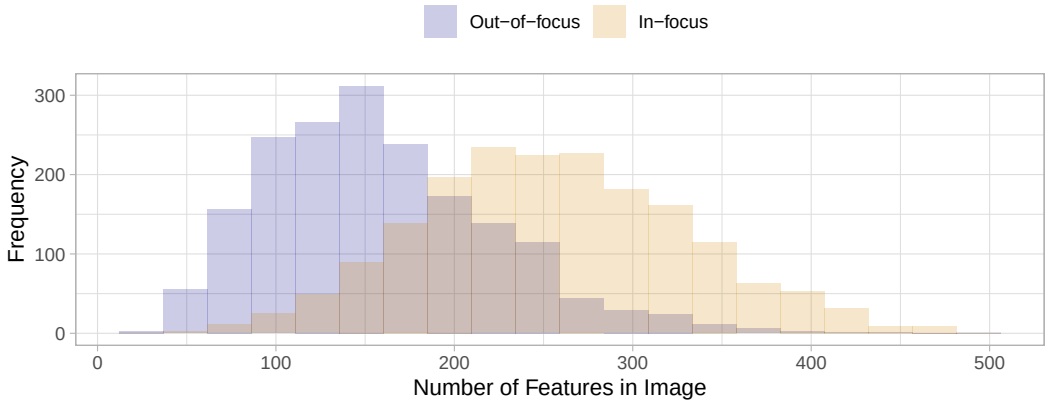


Fig. 8. The difference in the number of features found in blurry images (blur) compared to in-focus images (tan).

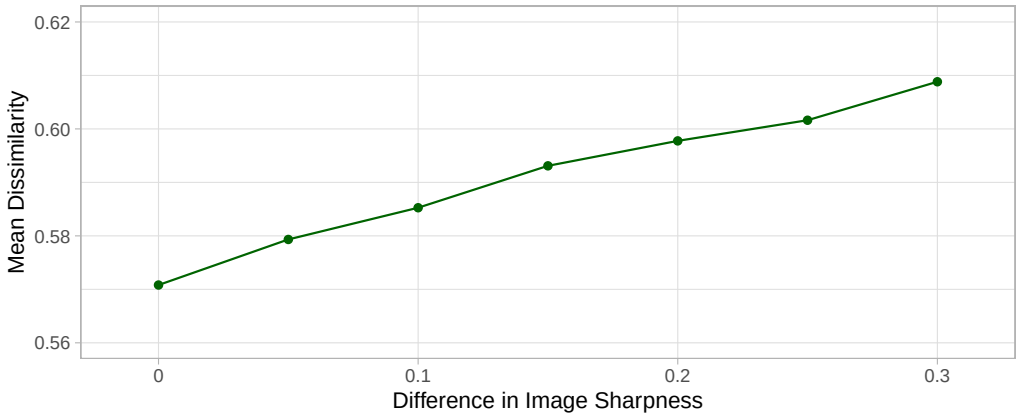


Fig. 9. The mean dissimilarity score between mated features as a function of the difference in sharpness (according to the S_3 metric) between the compared images.

5.3 Matching Accuracy

Several studies have explored the potential for using SIFT as well as related algorithms such as speeded up robust features (SURF) [53], and oriented FAST and rotated BRIEF (ORB) [54] to compare iris images. Results have been mixed, with some suggesting SURF is more

[3]Two feature points are *mated* if they represent the same feature in different images of the same iris.

accurate than SIFT [55, 56], others suggesting the reverse [57], and still others proposing a fusion of the two [58]. The primary goal of this paper is to explore the distinctiveness and permanence of local features, not to find the optimal feature-based approach to matching iris images. Nevertheless, one of the reasons SIFT was chosen over SURF is because its feature descriptors are designed to imitate the way a person's primary visual cortex is used to compare features of interest. Thus, it may be easier to integrate humans into the matching process when the SIFT algorithm is utilized.

5.3.1 Near-infrared Wavelengths

The similarity between iris images can be measured by the number of matched feature points between images (see Section 5). Figure 10 plots Detection Error Trade-off (DET) accuracy over the CMID dataset when using this metric. The figure demonstrates that about one in five nonmated comparisons still (erroneously) match at least two feature points. About 3 in 10 000 nonmated comparisons match 7 feature points. Fewer than one in a million nonmated comparisons match 13 or more feature points.

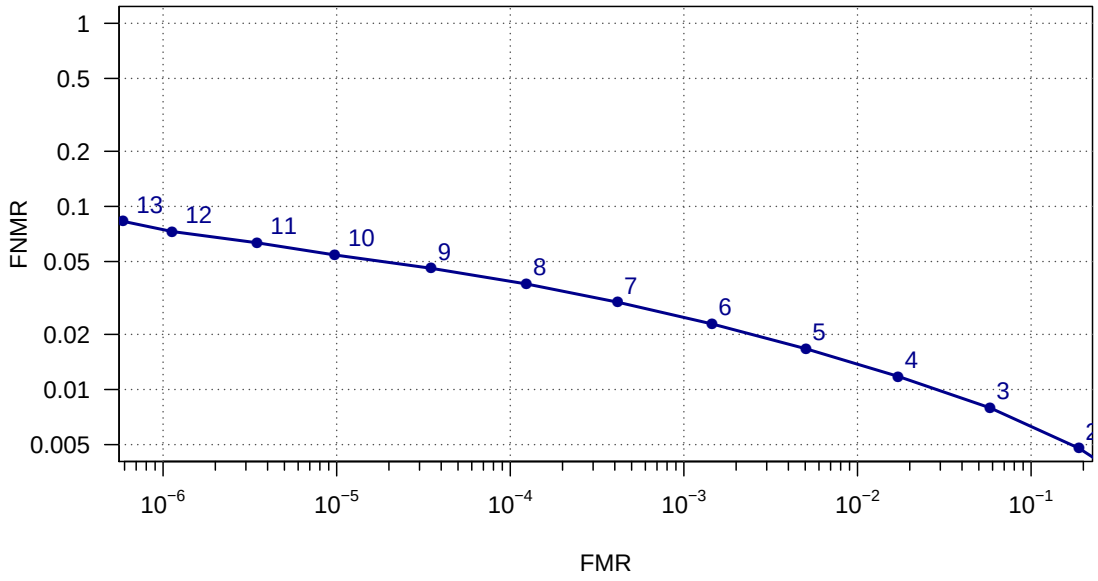


Fig. 10. DET accuracy of SIFT matcher, generated from 107 144 mated and 41 309 096 nonmated comparisons. The number printed alongside each point is the score, which represents the number of feature points could be paired between the compared iris images.

The accuracy of this DET cruve still lags behind the current state-of-the-art. In the NIST-administered IREX IX evaluation [59] (circa 2018), Decatur's matcher achieved a

False Non-match Rate (FNMR) of 0.0005 at an FMR of 10^{-4} over the same dataset compared to 0.038 for the current SIFT matcher.

5.3.2 Visible Wavelengths

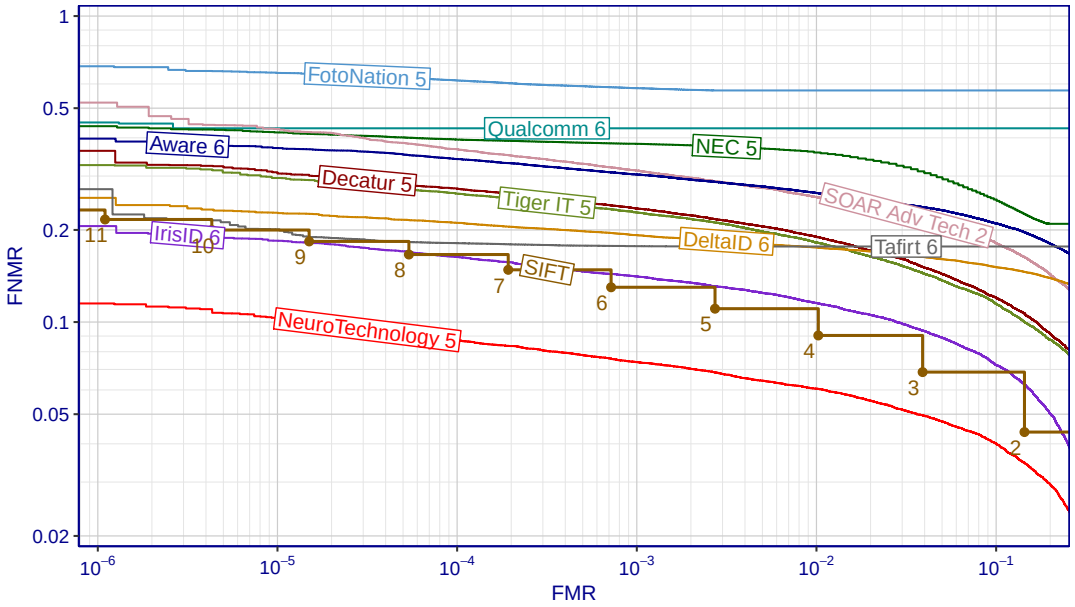


Fig. 11. DET accuracy for the SIFT matcher and other IREX 10 matchers operating over visible wavelength (620 nm) images. The number next to each point indicates the number of matched feature points.

Conventional iris recognition is performed over images acquired in the near infrared spectrum band (roughly 700 nm to 900 nm), but sometimes the images may have been acquired at visible wavelengths (400 nm to 700 nm). This is likely to be the case for many forensic applications. Figure 11 shows DET accuracy for images captured at 620 nm (corresponding to the color orange-red). The comparison score is the number of feature points that could be matched between the iris images (as in Section 5.3). Accuracy is poorer compared to images acquired at 800 nm, in large part because the melanin in the iris absorbs most light at visible wavelengths. At an FMR of 10^{-5}, the SIFT matcher achieves an FNMR of 0.20. In contrast, the most accurate IREX 9 matcher at visible wavelengths (submitted by NeuroTechnology) achieves an FNMR of 0.10 at the same FMR. The SIFT matcher nevertheless performs better than several of the commercial matchers submitted to IREX 9. This suggests local feature-based methods of matching may hold more promise than traditional methods at visible wavelengths. Further research is recommended.

5.4 Human-assisted Matching

One of the benefits of matching human-interpretable features is the potential to incorporate humans into the matching process. This can improve the robustness of matching decisions as well as make decisions more explainable. Automated matchers act naively at times, treating surface reflections, eyelashes, eyelid shadows, and specular highlights as though they are legitimate iris features. Despite the algorithms' difficulty, these false features are easily identifiable by human examiners.

Figure 12 shows a situation where the SIFT matcher fails to recognize that two of the four matched feature points are not, in fact, true iris features. The purple circles surround specular highlights that were falsely matched to one another. The yellow circles show another false match that occurs because the eyelash patterns create similar-appearing features. The red arrows show reflections off the cornea (a.k.a. Purkinji images) that appear similar in both images. Despite advancements in computer vision and object recognition, machine intelligence still lags behind human perception in many respects.

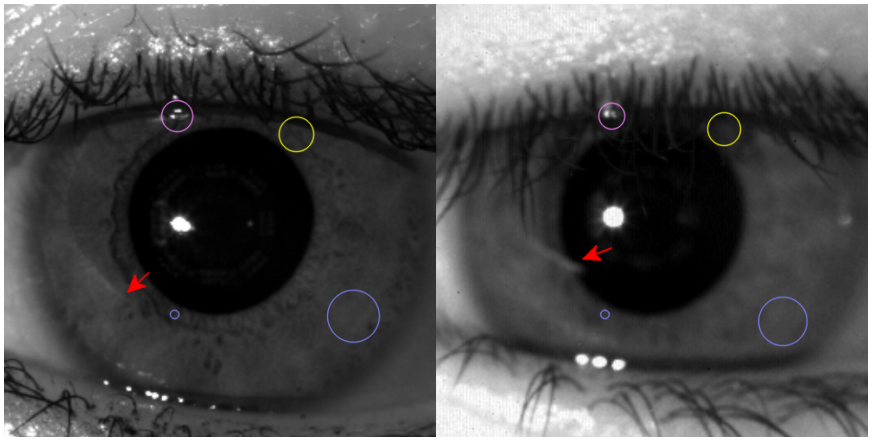


Fig. 12. Feature points paired between two different irides. The blue circles show pairings due to coincidental similarities. The purple and yellow circles show erroneous pairings introduced by specular highlights and eyelashes respectively.

5.5 Future Work

Matching visible features using algorithms other than SIFT should be investigated. SIFT was designed to perform 3D object recognition, a different problem than iris recognition. For this reason, some of the design decisions of SIFT do not carry over well. Feature point descriptors are scale invariant, but once the iris is localized its size is known, obviating the need for scale invariance. SIFT's approach to pairing features is also overly generalized for iris matching. SIFT searches for pairings across all locations, orientations, and scales, but the only significant degree of freedom when comparing iris images is rotation, which can often be determined manually with little effort. Similar logic holds for related approaches to 3D object recognition such as the speeded-up robust features (SURF) [60] algorithm. Convolutional Neural Networks have recently been applied to face recognition to great effect and may hold promise for iris matching.

6. Notre Dame Algorithm

Researchers at the University of Notre Dame developed an iris recognition algorithm that localizes and matches iris crypts and dark spots [14]. Their motivation was to devise a more visually interpretable approach to iris matching based on human perception and which they believed would likely be more defensible in law enforcement applications.

The Notre Dame feature extractor outputs a binarized representation of the original iris image. The core image processing operation is a top-hat transformation that retains smaller shapes that are darker than their surroundings. Shapes outside of a particular size range are disregarded, as are shapes covering areas where the variance in pixel intensity is too high for it to be a reliable representation of the iris texture at that location. Although the shapes found by the Notre Dame feature extractor algorithm often correspond to iris crypts, the matcher does not distinguish between crypts and other dark regions in the iris. Figure 13 shows the results of applying the Notre Dame feature extractor to a pseudo-polar representation of an iris image.

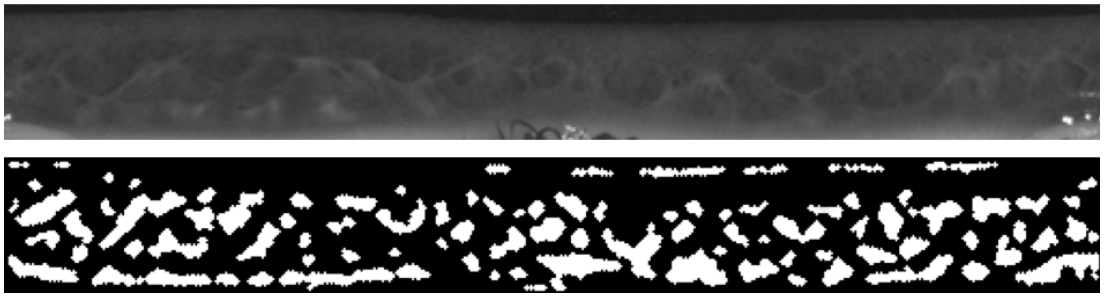
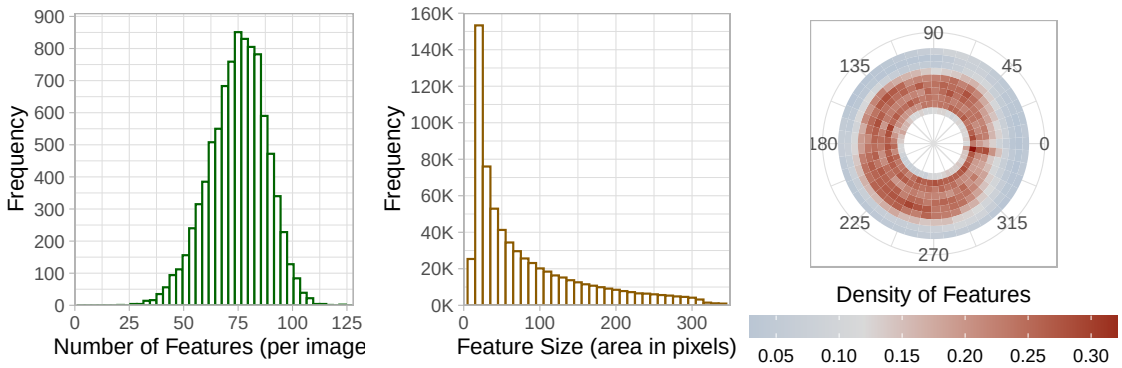


Fig. 13. An unwrapped pseudo-polar representation of an iris image (top image) along with corresponding shapes detected by Notre Dame's feature extractor (bottom image). The white shapes in the bottom image generally correspond to dark regions in the iris.



(a) Number of features found per iris image.

(b) Histogram of feature point sizes.

(c) Density plot showing where features are found.

Fig. 14. Basic statistical information for The University of Notre Dame's features.

6.1 Statistical Properties of Localized Features

Figure 14 presents basic statistics on the shapes localized by Notre Dame's feature extractor. Most iris images contain between roughly 50 and 100 detected features. The distribution of feature sizes skews toward smaller sizes, with half of the features having a pixel area between 15 and 50. About one in ten features have a size greater than 200 pixels. The feature extractor finds far more features in the pupillary ruff than in the outer ciliary zone. A light break appears at zero degrees in Figure 14c because the feature extractor does not properly handle the way pseudo-polar coordinates wrap around at zero degrees.

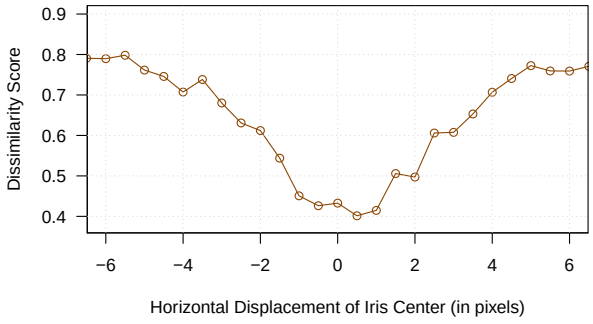


Fig. 15. Dissimilarity score for the ND Matcher when the pupil and iris centers for one of the compared images is horizontally shifted.

Notre Dame's matcher appears sensitive to variations in the specified locations of pupil and limbus boundaries. Figure 15 demonstrates this sensitivity for a particular comparison. Horizontally shifting the locations of the pupil and limbus centers by only 4 pixels increases the distance score from $\approx$0.4 to $\approx$0.7. The Notre Dame matcher requires accurate boundary localizations for optimal recognition accuracy. The manually specified boundary coordinates for the test dataset were sometimes a few pixels off, which would detrimentally impact accuracy for Notre Dame's matcher.

6.2 Future Work

The Notre Dame feature detector is a good starting point. It searches for dark spots such as crypts, but not all irides have crypts and other types of features exist. The feature detector could be expanded to detect other distinctive features such as localized small diameter pigmented lesions that can be freckles, nevi, or other lesions. The current implementation could be modified to properly handle wrap-arounds at zero degrees in the pseudo-polar images. The matcher could be made more robust to inaccuracies in boundary localization or manual tweaking of their locations could be performed.

References

[1] J. Daugman, "Iris recognition," *American scientist*, vol. 89, no. 4, pp. 326–333, 2001.

[2] J. Daugman, "How Iris Recognition Works," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 1, pp. 21–30, 2004.

[3] J. Daugman, "Information Theory and the IrisCode," *IEEE Transactions on Information Forensics and Security*, 2015. Preprint.

[4] J. Daugman, "Evolving Methods in Iris Recognition," in *IEEE International Conference on Biometrics: Theory, Applications, and Systems,(BTAS07),(online). http://www.cse.nd.edu/BTAS_07/John_Daugman_BTAS.pdf*, pp. 1–105, 2016. Accessed: 2021-02-23.

[5] J. Daugman and C. Downing, "Searching for Doppelgängers: Assessing the Universality of the IrisCode Impostors Distribution, volume = 5, year = 2016,," *IET Biometrics*, no. 2, pp. 65–75.

[6] L. Masek, *Recognition of human iris patterns for biometric identification*. Bachelor's thesis, University of Western Australia, 2003.

[7] H. R. E. Wallis, "The significance of Brushfield's spots in the diagnosis of mongolism in infancy," *Archives of disease in childhood*, vol. 26, pp. 495–500, 12 1951.

[8] "Human eye." https://www.britannica.com/science/human-eye. Accessed: 2021-03-21.

[9] E. Wolfflin *et al.*, "Ein klinischer beitrag zur kenntniss der structur der iris," *Arch Augenheilkd*, vol. 45, pp. 1–4, 1902.

[10] D. Gold and R. Lewis in *Clinical Eye Atlas*, pp. 596–599, Elsevier, 2002.

[11] M. J. Tarr, "Visual pattern recognition," *Encyclopedia of psychology*, pp. 1–4, 2000.

[12] N. R. Council, *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: The National Academies Press, 2009.

[13] D. G. Lowe, "Object Recognition from Local Scale-Invariant Features," in *Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2*, ICCV '99, (Washington, DC, USA), pp. 1150–, IEEE Computer Society, 1999.

[14] J. Chen, F. Shen, D. Z. Chen, and P. J. Flynn, "Iris Recognition Based on Human-Interpretable Features," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 1476–1485, July 2016.

[15] A. Biedermann, S. Bozza, F. Taroni, and J. Vuille, "Are inconclusive decisions in forensic science as deficient as they are said to be?," *Frontiers in psychology*, vol. 10, p. 520, 2019.

[16] G. W. Quinn, P. Grother, and J. Matey, "IREX IX Part One: Performance of Iris Recognition Algorithms." https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8207.pdf, 2018.

[17] P. Grother, G. Quinn, J. Matey, M. Ngan, W. Salamon, G. Fiumara, and C. Watson, "IREX III: Performance of Iris Identification Algorithms." https://www.nist.gov/itl/iad/image-group/irex-iii-homepage, 2011.

[18] G. Quinn and P. Grother, "IREX III Supplement I: Failure Analysis." https://www.nist.gov/itl/iad/image-group/irex-iii-homepage, 2011.
[19] J. Daugman, "New methods in iris recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, pp. 1167–1175, Oct 2007.
[20] S. C. Dass, S. Pankanti, S. Prabhakar, and Y. Zhu, *Individuality of Fingerprints*, pp. 741–751. Boston, MA: Springer US, 2009.
[21] A. Ross, S. C. Dass, and A. K. Jain, "Fingerprint warping using ridge curve correspondences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 19–30, 2005.
[22] J. Daugman, "The importance of being random: statistical principles of iris recognition," *Pattern recognition*, vol. 36, no. 2, pp. 279–291, 2003.
[23] G. W. Quinn, P. Grother, and M. Ngan, "IREX IV Part 1: Evaluation of Iris Identification Algorithms." https://www.nist.gov/publications/irex-iv-part-1-evaluation-iris-identification-algorithms, 2014.
[24] C. Schwab, C. Mayer, I. Zalaudek, R. Riedl, M. Richtig, W. Wackernagel, R. Hofmann-Wellenhof, G. Richtig, G. Langmann, L. Tarmann, *et al.*, "Iris freckles a potential biomarker for chronic sun damage," *Investigative ophthalmology & visual science*, vol. 58, no. 6, pp. BIO174–BIO179, 2017.
[25] E. T. G. W. Quinn, J. R. Matey and P. J. Grother, "IREX V: Guidance for Iris Image Collection." https://www.nist.gov/itl/iad/image-group/irex-v-homepage, 2014.
[26] "MedlinePlus: coloboma." https://medlineplus.gov/genetics/condition/coloboma. Accessed: 2021-02-23.
[27] C. L. Wilkerson, N. A. Syed, M. R. Fisher, N. L. Robinson, D. M. Albert, *et al.*, "Melanocytes and iris color: light microscopic findings," *Archives of Ophthalmology*, vol. 114, no. 4, pp. 437–442, 1996.
[28] M. Edwards, D. Cha, S. Krithika, M. Johnson, and E. J. Parra, "Analysis of iris surface features in populations of diverse ancestry," *Royal Society open science*, vol. 3, no. 1, p. 150424, 2016.
[29] A. E. Christensen, T. Fiskerstrand, P. M. Knappskog, H. Boman, and E. Rødahl, "A novel adamtsl4 mutation in autosomal recessive ectopia lentis et pupillae," *Investigative ophthalmology & visual science*, vol. 51, no. 12, pp. 6369–6373, 2010.
[30] "American Academy of Ophthalmology: nevus." https://www.aao.org/eye-health/diseases/what-is-nevus.
[31] "American Academy of Ophthalmology: iris hypoplasia." https://www.aao.org/eye-health/ask-ophthalmologist-q/what-is-iris-hypoplasia. Accessed: 2021-02-23.
[32] S. Shareef, "Laser peripheral iridoplasty," in *Operative Dictations in Ophthalmology*, pp. 177–180, Springer, 2017.
[33] M. Millodot, *Dictionary of Optometry and Visual Science E-Book*. Elsevier Health Sciences, 2014.
[34] "Farlex partner medical dictionary: Pupillary ruff." https://medical-dictionary.thefreedictionary.com/pupillary+ruff. Accessed: 2021-02-23.
[35] H. Gray, *Gray's anatomy*. Arcturus Publishing, 2009.

[36] D. Etter, J. Webb, and J. Howard, "Collecting Large Biometric Datasets: A Case Study in Applying Software Best Practices," *CrossTalk: The Immutable Laws of Software Development*, pp. 4–8, 2014.
[37] U. Park, S. Pankanti, and A. K. Jain, "Fingerprint verification using sift features," in *Biometric Technology for Human Identification V*, vol. 6944, p. 69440K, International Society for Optics and Photonics, 2008.
[38] D. R. Kisku, A. Rattani, E. Grosso, and M. Tistarelli, "Face identification by SIFT-based complete graph topology," in *2007 IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 63–68, 2007.
[39] J. Chen and Y. Moon, "Using SIFT features in palmprint authentication," Dec 2008.
[40] J. D. Bustard and M. W. Nixon, "Robust 2D Ear Registration and Recognition Based on SIFT Point Matching," *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pp. 1–6, 2008.
[41] P.-O. Ladoux, C. Rosenberger, and B. Dorizzi, "Palm Vein Verification System Based on SIFT Matching," in *Proceedings of the Third International Conference on Advances in Biometrics*, ICB '09, (Berlin, Heidelberg), pp. 1290–1298, Springer-Verlag, 2009.
[42] C. Belcher and Y. Du, "Region-based SIFT approach to iris recognition," *Optics and Lasers in Engineering*, vol. 47, pp. 139–147, 01 2009.
[43] F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, and J. Ortega-Garcia, "Iris recognition based on SIFT features," in *2009 First IEEE International Conference on Biometrics, Identity and Security (BIdS)*, pp. 1–8, Sep. 2009.
[44] H. Mehrotra, B. Majhi, and P. Gupta, "Robust iris indexing scheme using geometric hashing of SIFT keypoints," *Journal of Network and Computer Applications*, vol. 33, pp. 300–313, 05 2010.
[45] M. S. Sunder and A. Ross, "Iris Image Retrieval Based on Macro-features," in *2010 20th International Conference on Pattern Recognition*, pp. 1318–1321, Aug 2010.
[46] S. Sun, S. Yang, and L. Zhao, "Noncooperative bovine iris recognition via sift.," *Neurocomputing*, vol. 120, pp. 310–317, 2013.
[47] R. Zhu, J. Yang, and R. Wu, "Iris recognition based on local feature point matching," Oct 2006.
[48] E. W. Weisstein, "Principal Curvatures." http://mathworld.wolfram.com/PrincipalCurvatures.html, 2019.
[49] S. Edelman, N. Intrator, and T. Poggio, "Complex cells and object recognition," *Neural Information Processing Systems (NIPS), Section: Visual Processing*, 06 1997.
[50] M. S. Sunder and A. Ross, "Iris image retrieval based on macro-features," in *Proceedings of the 2010 20th International Conference on Pattern Recognition*, pp. 1318–1321, IEEE Computer Society, 2010.
[51] J. Geng, Y. Li, and T. Chian, "Sift based iris feature extraction and matching," in *Geoinformatics 2007: Geospatial Information Science*, vol. 6753, p. 67532F, International Society for Optics and Photonics, 2007.
[52] C. T. Vu and D. M. Chandler, "S3: A spectral and spatial sharpness measure," in *2009*

First International Conference on Advances in Multimedia, pp. 37–43, July 2009.

[53] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision – ECCV 2006* (A. Leonardis, H. Bischof, and A. Pinz, eds.), (Berlin, Heidelberg), pp. 404–417, Springer Berlin Heidelberg, 2006.

[54] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *2011 International Conference on Computer Vision*, pp. 2564–2571, 2011.

[55] A. Ignat and I. Păvăloi, "Occluded iris recognition using surf features," in *16th International Conference on Computer Vision Theory and Applications*, 2021.

[56] H. Mehrotra, P. K. Sa, and B. Majhi, "Fast segmentation and adaptive SURF descriptor for iris recognition," *Mathematical and Computer Modelling*, vol. 58, no. 1, pp. 132–146, 2013. Financial IT and Security and 2010 International Symposium on Computational Electronics.

[57] M. Kamal Majeed, "Surf and sift descriptors using wavelet transforms for iris recognition," *Surf and Sift Descriptors Using Wavelet Transforms for Iris Recognition*, vol. 13, no. 4, pp. 2361–2373, 2020.

[58] S. Bakshi, S. Das, H. Mehrotra, and P. K. Sa, "Score level fusion of sift and surf for iris," in *2012 International Conference on Devices, Circuits and Systems (ICDCS)*, pp. 527–531, 2012.

[59] G. W. Quinn, P. Grother, and J. Matey, "IREX IX Part II: Multispectral Iris Recognition." https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8252.pdf, 2019.

[60] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

A. SIFT Parameters

The parameters in the table below were used to extract SIFT feature points in the iris images used in this study.

Number of Octal Layers	3
Peak Threshold	0.04
Edge Threshold	10