NIST Technical Note 2097

Analysis of Iris Images from Twins Day – 2010

James R. Matey
Patrick J. Grother
George W. Quinn

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2097

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

NIST Technical Note 2097

Analysis of Iris Images from Twins Day – 2010

James R. Matey
Patrick J. Grother
George W. Quinn
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2097

June 2020

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

All research reported in this paper based on data obtained from human subjects was carried out under protocols reviewed and approved by the NIST Research Protection Office (RPO) and/or the NIST Institutional Review Board. The NIST protocols relevant for this paper are:

- ITL-2019-0142 Evaluation of Biometric Algorithms on Twins

All images shown in this paper have been checked for conformance to protocols/licensing regarding publication.

The twins data used in this paper was originally collected by J. M. Dawson and colleagues, as described in the paper by Sabatier et al.[1]. The data transfer agreement covering the transfer of this data to NIST does not permit redistribution of the data.

National Institute of Standards and Technology Technical Note 2097
Natl. Inst. Stand. Technol. Tech. Note 2097, 15 pages (June 2020)
CODEN: NTNOEF

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2097

Abstract

Each year since 1976, the town of Twinsburg, Ohio has held a *Twins Days Festival*[1]. Over the past 43 years over $77,000$ sets of twins and multiples have attended – with many repeat visits over the years. Prof. Jeremy Dawson and other staff from the the West Virginia University Biometrics Collection Lab have engaged with the Twins Days organizers since 2010 to collect biometric data from twins attending the festivals; the collections were carried out under Institutional Review Board (IRB) protocols and with informed consent of the individuals participating[1]. NIST recently obtained copies of the Twins Day data for use in ongoing evaluations of face, finger and iris recognition algorithms. This paper is an initial analysis of the iris images taken at the 2010 festival.

We see no evidence that the non-mated distribution for identical twin pairs is significantly different from that for unrelated individuals for two commercial iris recognition algorithms. This is not a surprise; Daugman demonstrated, using iris2pi, that a non-mated distribution of match scores from genetically identical left/right pairs from the same individual is not significantly different from that for pairs from different individuals[2] and recently demonstrated results on twins similar to what we found[3].

Analysis of iris images from subsequent festivals as well as face and finger images are planned.

Key words

iris recognition; biometrics; twins.

[1]See https://twinsdays.org/

Table of Contents

List of Tables

List of Figures

1. Introduction

Each year since 1976, the town of Twinsburg, Ohio has held a *Twins Days Festival*[2]. Over the past 43 years over $77,000$ sets of twins and multiples have attended – with many repeat visits over the years. Prof. Jeremy Dawson and other staff from West Virginia University Biometrics Collection Lab have engaged with the Twins Days organizers since 2010 to collect biometric data from twins attending the festivals; the collections were carried out under Institutional Review Board (IRB) protocols and with informed consent of the individuals participating[1]. NIST recently obtained copies of the Twins Day data for use in ongoing evaluations of face, finger and iris recognition algorithms.

We note that NIST does not have any redistribution rights for the twins data – we cannot provide copies of the data to other researchers. Interested readers may contact the authors to make suggestions about additional analyses that might be useful to the biometric community.

This paper presents a preliminary analysis of the iris images taken at the 2010 festival. Analysis of iris images from subsequent festivals as well as analysis of face and finger data are planned. We note that a recent paper by Daugman[3] on the topic of twins and and iris recognition became available while this paper was in review.

To understand the discussion below, the reader should be aware of the idea of mirror twins: otherwise identical twins which display reversed asymmetry such as opposite handedness – one of the twin pairs is right handed, the other left[4] [5].

2. Description of Analysis

The twins iris images were collected using commercial off the shelf iris cameras; the 2010 collection employed an Oki IrisPass-M device[6][3] and the NIST MBARK software platform[7] . The IrisPass-M is a two eye camera that captures both left and right eyes with near simultaneity. For each capture-event, the collection system captured two images of each eye resulting in four iris images per capture-event.

For the 2010 Twins Day Festival all subjects, except one, participated in a single capture event resulting in four iris images; one subject (a twin) participated in two capture events resulting in eight images for that subject. Table 2 summarizes the 2010 collection. Table 3 summarizes the match pairs in this analysis of that collection. We note that the images in the mated pairs were taken in close temporal proximity (with the exception of the single subject who engaged in two capture-events); though they differ in the light direction, this likely makes the match scores for the mated pairs better than might be expected in an operational setting as discussed in the ISO-IEC 19795-1 biometric testing standard [8]: *"for some modalities, performance a short time after enrollment, when the user appearance and behavior has changed very little, is far better than that obtained weeks or months later."*

The iris images were then processed in pairs to generate match scores using a commercial version of the iris2pi iris recognition algorithm whose internals have been described in detail by Daugman[9] [10] [11] [12] and variants of which have been implemented by academics, see for example Masek[13].

We also used another commercial iris recognition algorithm, Neurotechnology/VeriEye Version 10, for which the internals are not well known.

[2]See https://twinsdays.org/

[3]This device is no longer in production.

Table 1. Meta-data Variable/Value Definitions

Variable/Value	**Description**
SubjectID	coded subject identifier
SubjectEye	left/right
Mated	$(SubjectID_1 == SubjectID_2)\, AND\, (SubjectEye_1 == SubjectEye_2)$
Non-mated	NOT Mated
Identical	$Subject_1\, AND\, Subject_2$ both claim the other as identical twins
Identical-Mirror	$Subject_1\, AND\, Subject_2$ both claim the other as identical mirror twins
Identical-Ambiguous	$Subject_1\, AND\, Subject_2$ both claim the other as twins, but disagree if they are mirror
Non-mated-Identical	Non-mated AND Identical
Non-mated-None	Non-mated AND NOT (Identical OR Identical-Mirror OR Identical-Ambiguous)

Table 2. Summary of 2010 Twins Day Iris Image Collection

Number of iris images	52
Number of subjects	12
Number iris images/subject/capture-event	4
Number of identical twin images	44
Number of mirror twin images	8
Number of successful template operations (iris2pi)	52
Number of successful template operations (VeriEye)	52

Table 3. Summary of Match Pairs for Analysis of 2010 Twins Day Iris Images

Number of mated pairs	34
Number of non-mated pairs	1292
Number of non-mated identical twin pairs	80
Number of non-mated mirror twin pairs	16
Number of non-mated non-twin pairs	1196

The resulting match scores and the non-PII (personally identifiable information) meta-data were then read into data-frames in R[14] and the meta-data was joined with the match scores on the basis of the coded ID's that labeled both the images and the meta-data. The meta-data variables of interest in this paper are defined in Table 1. We note that the Non-Mated-None category includes comparisons between the left and right eyes of the same individual; Daugman previously demonstrated that, for iris2pi, the left/right match scores for the same subjects have the same distribution as those between un-related subjects[2].

In our review of the entire Twins Days dataset (2010 to 2019), we found occasional examples of ambiguities in the meta-data, as is common in most data collections, particularly for values that are self reported and/or entered by hand. The two types of ambiguities that we encountered are:

- Inconsistent relationship status: The meta-data for the Twins Day collections includes a self assessment of whether a twin pair is identical or identical mirror[4] [5]. In some cases that assessment is not consistent between the individuals of a twin pair. One says they are twins, the other that they are mirror twins.
- Inconsistent twin pairing: An individual listed as a twin, linked to more than one individual. Such cases likely indicate an entry error at the time of collection.

We found no twin pairing inconsistencies nor mirror twin ambiguities in the 2010 data. For this analysis of the 2010 data we dropped the mirror twin data (one twin pair) from the non-mated twin distribution analysis because one pair is not enough data to permit reliable tests of a mirror twin effect. We will examine the effects of mirror twins in subsequent analyses of the full data set.

We used R[14] to produce plots and statistical tests on the match score data:

- Box plots[15] of match scores by algorithm and mated/non-mated
- Box plots of non-mated match scores by algorithm and relationship
- Kolmogorov-Smirnov tests[16] comparing the distributions of non-mated scores between identical twins and no relationship
- Quantile-Quantile plots[17] comparing the distributions of non-mated scores between identical twins and no relationship

In the plots below, the match scores are raw scores as reported by the algorithms. The iris2pi scores are Hamming distances which are dissimilarity scores. The VeriEye scores are proprietary similarity scores.

3. Results

In the discussion below, the number of non-mated identical twins scores per algorithm is 80; the number of non-mated, not-between-twins scores per algorithm is 1196; the number of mated pairs is 34 per algorithm.

Figure 1 compares the mated and non-mated scores by algorithm without regard to relationship. For both algorithms we see good separation between the mated and non-mated categories. The scores here are consistent with absence of labeling errors for subject ID and eye.

In subsequent plots we consider the effect of relationship on the non-mated distribution. The relationships between individuals fall into two categories:

- No relation (none): the score is between subjects who are not twins; including left/right comparisons for a single subject which have been demonstrated to follow the same distribution as comparisons between un-related subjects[2].
- Identical: the score is between identical twins; this includes L/L, L/R and R/L comparisons between the two individuals.

The Kolmogorov-Smirnov tests in table 4 show that the differences between the non-mated distributions (raw scores) for subjects who are not related are not significantly different from those for identical twins, as suggested by the box plots in figure 2 and the QQ plots in figure 3.

Table 4. Kolmogorov-Smirnov tests of whether distributions are different

Distributions	Algorithm	KS p value
No relationship/identical twins	iris2pi	0.12
No relationship/identical twins	VeriEye	0.84

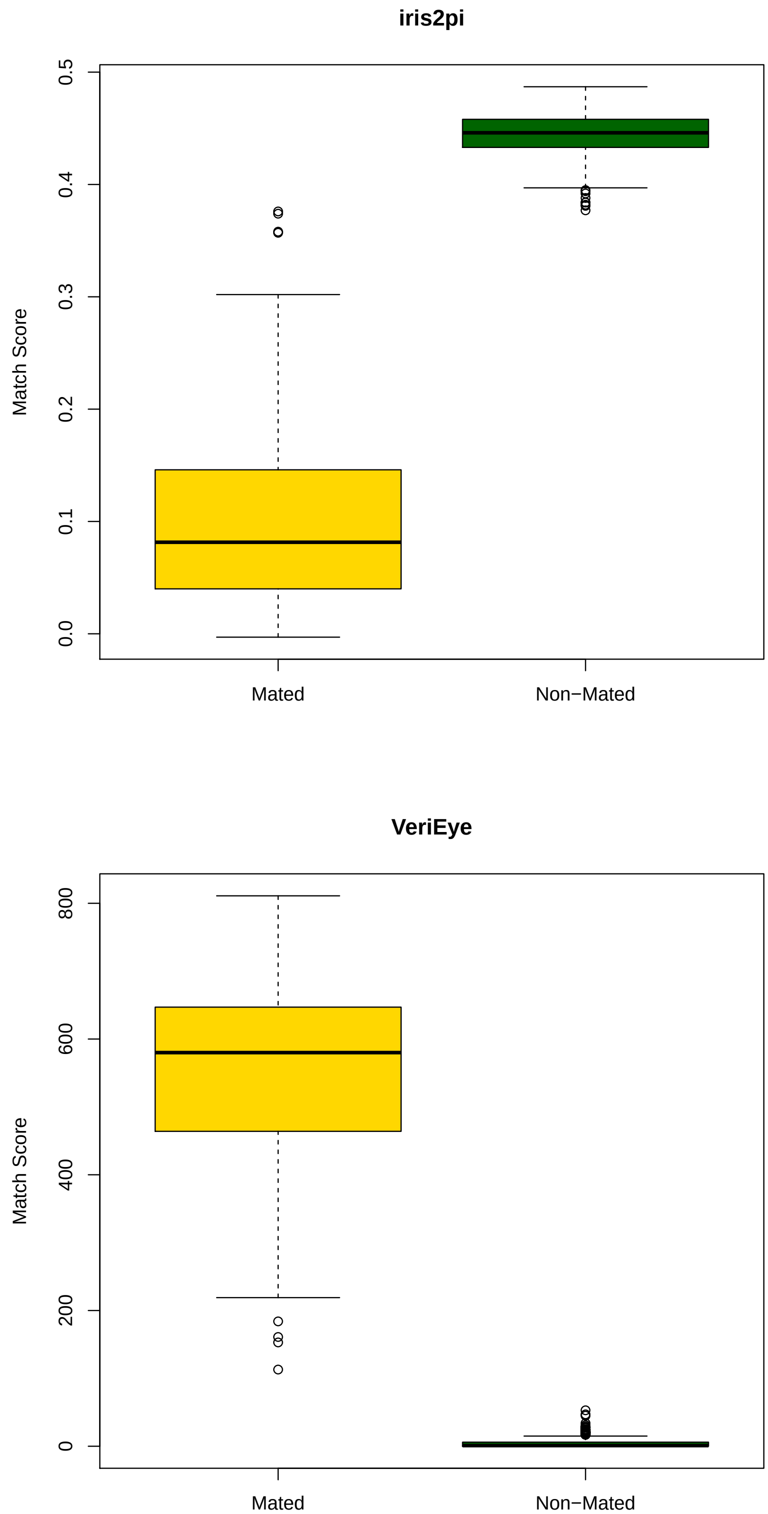


Fig. 1. Box plots of mated and non-mated raw match scores by algorithm. There is excellent separation of the mated and non-mated scores for both algorithms. The mated score performance is likely better than would be expected in an operational setting because the mated pairs were taken in close temporal proximity.

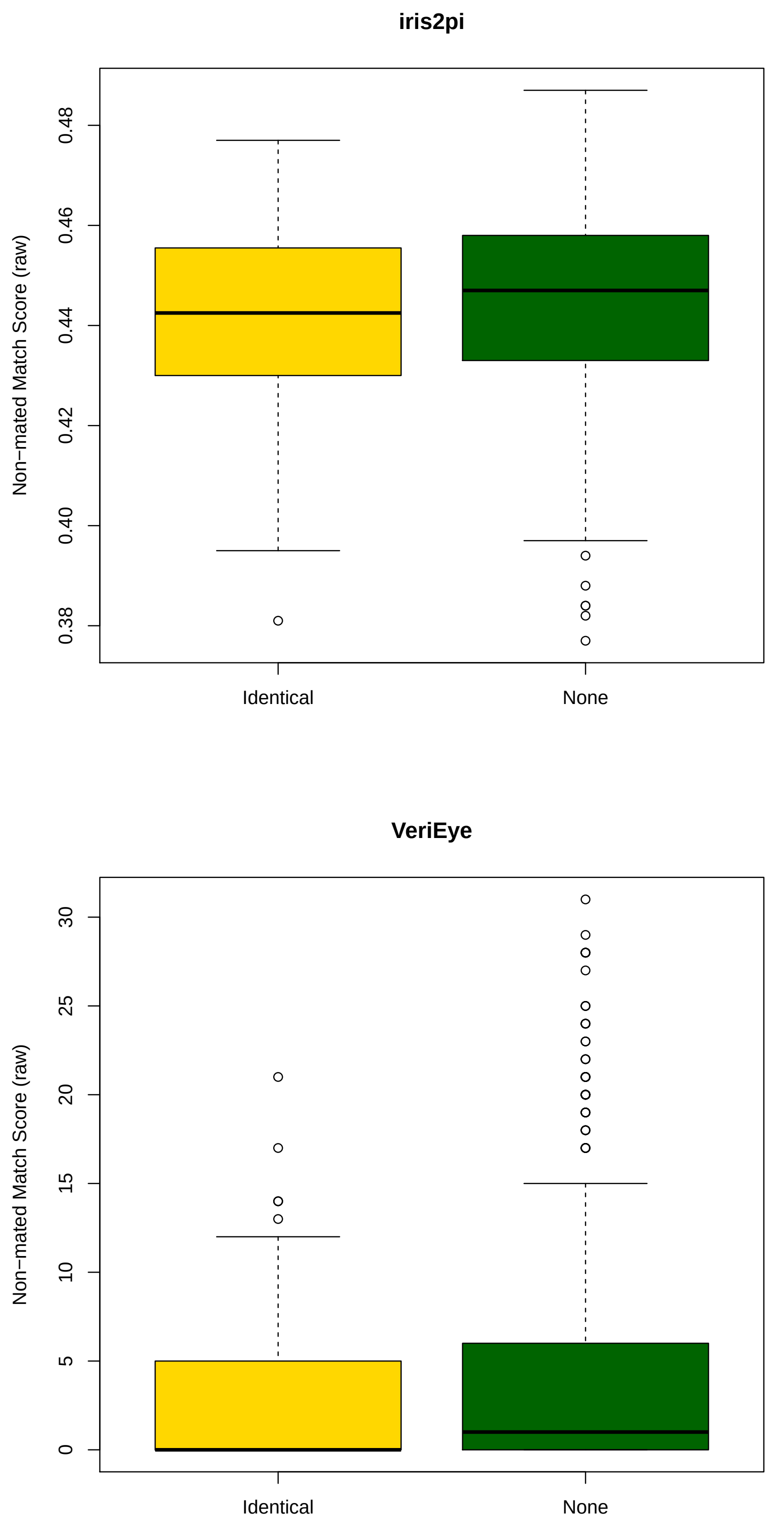
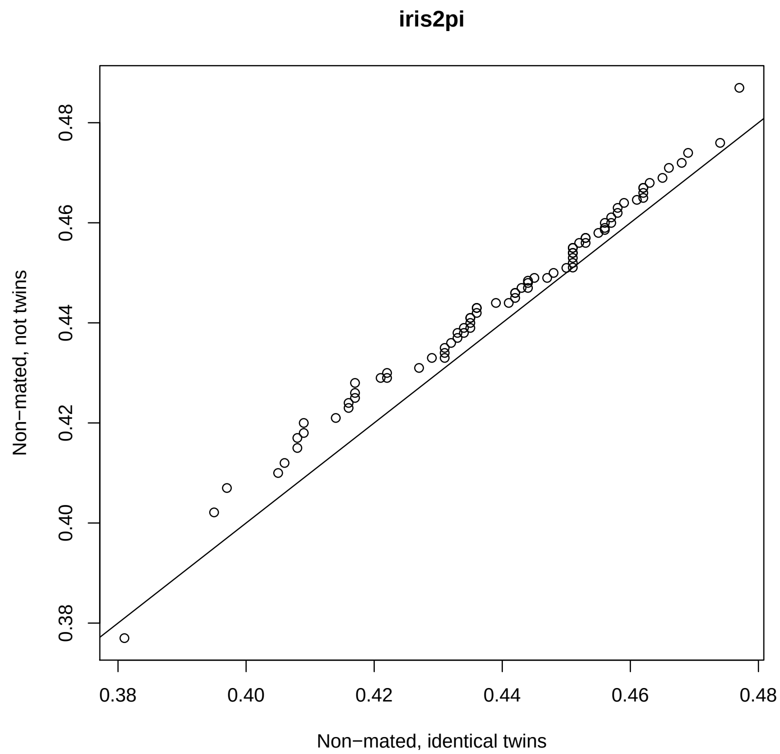


Fig. 2. Box plots of non-mated raw match scores by algorithm and relationship. The box plots for both algorithms show little difference between the two relationship cases.



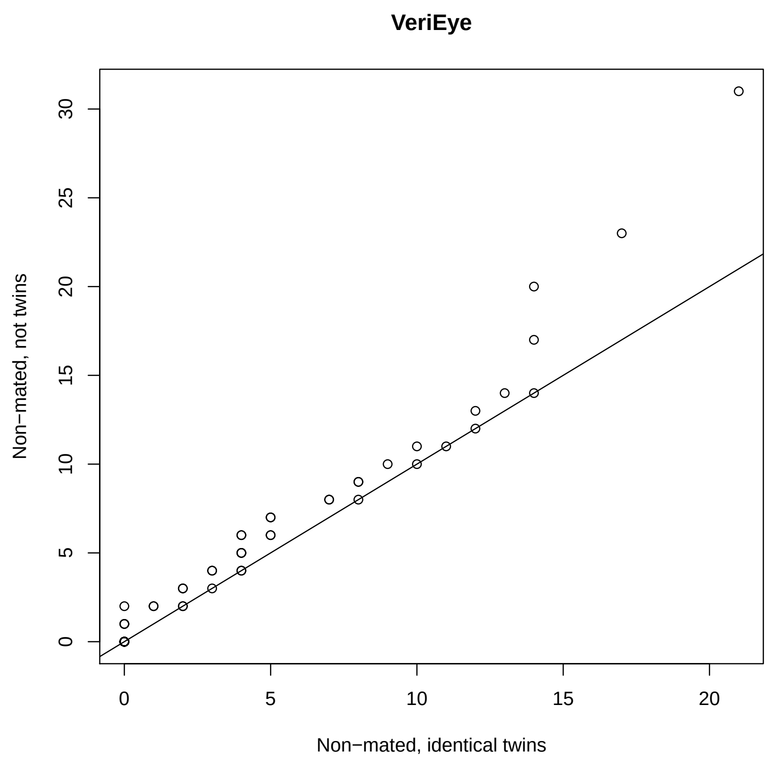


Fig. 3. QQ plots of raw non-mated scores, identical vs. no relationship, for iris2pi and VeriEye. The plots closely follow the diagonal which would be expected if the distributions are the same except for some outliers on the right of the VeriEye plot. The outliers suggest there are a few high VeriEye similarity scores in the "none" category relative to the "identical" category, consistent with the boxplot in figure 2

4. Conclusions and Future Plans

In this preliminary analysis, we applied two commercial iris algorithms to iris images collected from twins at Twins Day. We see no evidence that the resulting non-mated distributions for iris recognition for identical twin pairs are significantly different from that for unrelated individuals. This is not a surprise; Daugman demonstrated that a non-mated distribution of match scores from genetically identical left/right pairs from the same individual is not significantly different from that for pairs from different individuals[2]. That paper also found no significant difference for a small number (six pair-wise comparisons) of monozygotic twins. Recently, while this paper was in preparation, Daugman presented new results[3] that are largely in agreement with the results presented herein, but which show a quite small (HD difference of 0.003) but statistically significant difference between the unmated distributions for twins and un-related persons in a particular database. Daugman argues that the effect could be genetic or environmental.

Our future plans include analysis of subsequent iris collections and analysis of the face and finger data collected at Twins Day Festivals as well as incorporation of the twins datasets into ongoing biometric evaluations at NIST. Those analyses may help resolve the genetic/environmental question posed by Daugman.

Acknowledgments

Our thanks to Prof. Jeremy Dawson, the staff of the West Virginia University Biometrics Collection Lab, and their sponsors for collecting the data, making it available to NIST, and for helpful discussions regarding details of the data collection and the associated meta-data.

Our thanks to both Dr. David Ackerman and Kayee Hanaoka for thoughtful comments that improved the clarity and readability of the paper.

Our thanks to Prof. John Daugman for providing a preprint of his recent work[3] in this area.

References

[1] Sabatier SB, Trester MR, Dawson JM (2019) Measurement of the impact of identical twin voices on automatic speaker recognition. *Measurement* 134:385–389.

[2] Daugman J, Downing C (2001) Epigenetic randomness, complexity and singularity of human iris patterns. *Proceedings of the Royal Society of London Series B: Biological Sciences* 268(1477):1737â€“1740. URL http://rspb.royalsocietypublishing.org/content/268/1477/1737.short.

[3] Daugman J, Downing C (2020) Broken symmetries, random morphogenesis, and biometric distance. *IEEE Trans on Biometrics, Behavior, and Identity Science* .

[4] Twins Research Australia (2020) Mirror-image twins, . URL www.twins.org.au/twins-and-families/about-twins/171-mirror-image-twins.

[5] Newman H (1928) Studies of human twins: II. Asymmetry reversal, of mirror imaging in identical twins. *The Biological Bulletin* 55(4):298–315.

[6] Itoda S, Kikuchi M, Wada S (2006) Fully automated image capturing-type iris recognition device ”IrisPass-M”. *OKI Technology Review* 73.

[7] Micheals RJ, Mangold K, Hanaoka K, Marshall K (2009) Multimodal Biometric Application Resource Kit (MBARK), . URL www.nist.gov/services-resources/software/multimodal-biometric-application-resource-kit-mbark.

[8] JTC1, Working Group 5 (2005) ISO/IEC 19795-1 Biometric Performance Testing and Reporting: Principles and Framework, . URL webstore.ansi.org.

[9] Daugman J (2001) Iris recognition. *American scientist* 89(4):326–333. URL https_//www.americanscientist.org/issues/issue.aspx?id=737&y=2001&no=4&content=true&page=4&css=print.

[10] Daugman J (2004) How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on* 14(1):21–30. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1262028.

[11] Daugman J (2006) Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE* 94(11):1927–1935.

[12] Daugman J (2016) Evolving methods in iris recognition. *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, . URL http://www.cse.nd.edu/BTAS_07/John_Daugman_BTAS.pdf.

[13] Masek L (2003) *Recognition of human iris patterns for biometric identification*. Ph.D. thesis. Master’s Thesis, Univesity of Western Australia, .

[14] R Core Team (2019) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing Vienna, Austria, . URL https://www.R-project.org/.

[15] Wickham H, Stryjewski L (2011) 40 years of boxplots. *Am Statistician* URL vita.had.co.nz/papers/boxplots.pdf.

[16] Arnold TB, Emerson JW (2011) Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal* 3(2). URL pdfs.semanticscholar.org/d9c0/69f62de3dc1d13c76fe8e38cffff248fdb6b.pdf#page=34.

[17] Marden JI, et al. (2004) Positions and qq plots. *Statistical Science* 19(4):606–614.