

NIST Technical Note 2155

**The NIST IrisDaily Dataset:
Description and Initial Analysis**

James R. Matey

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2155>

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

NIST Technical Note 2155

The NIST IrisDaily Dataset: Description and Initial Analysis

James R. Matey
*Information Access Division/Image Group
Information Technology Laboratory*

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2155>

May 2021



U.S. Department of Commerce
Gina M. 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 & Director, National Institute of Standards and Technology*

Certain commercial entities, equipment, software 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, software or equipment are necessarily the best available for the purpose.

NIST policy requires that all research is reviewed and approved by the human subjects and/or vertebrate animal review process before research involving human or animal subjects is conducted. The National Institute of Standards and Technology Institutional Review Board reviewed and approved the protocol for this project and all subjects provided informed consent in accordance with 15 CFR 27, the Common Rule for the Protection of Human Subjects. The NIST protocols relevant for this paper are:

- ITL-16-0002: Data Collection in Support of Evaluation of Biometric Algorithms

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

National Institute of Standards and Technology Technical Note 2155
Natl. Inst. Stand. Technol. Tech. Note 2155, 13 pages (May 2021)
CODEN: NTNOEF

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2155>

Abstract

This paper describes a dataset, IrisDaily-NIST, that is under construction by NIST. The purpose of the dataset is to enable evaluation of long and short term variability in comparison scores under optimal conditions. At the present, this dataset has (nearly) daily iris images of a single subject collected with a commercial iris camera over a period of approximately 4 months.

We present an initial analysis of the data using a commercially available algorithm.

Key words

iris recognition; biometrics; match score variation; comparison score variation.

Table of Contents

1	Introduction	1
2	Description of Data and Analysis	1
3	Results	3
4	Future Plans	7

List of Tables

Table 1	Description of IrisDaily-NIST Dataset	2
---------	---------------------------------------	---

List of Figures

Fig. 1	Example of an image from the IrisDaily dataset.	2
Fig. 2	Box-plots of comparison scores for mated and non-mated image pairs.	4
Fig. 3	Box-plots and histograms of comparison scores for mated and non-mated image pairs.	4
Fig. 4	Box-plots of iris2pi comparison scores by time delta	5
Fig. 5	QQ plots of comparison scores for mated and non-mated image pairs.	6

1. Introduction

In the deployment of biometric identification systems for border crossing and access control, an important question is the long term viability of system enrollments. Responding to such questions requires an understanding of long term variability of the biometric modality over the population of users. There has been significant work on long term variation in iris recognition for large populations, most notably the IREX VI studies[1].

In forensic settings[2], a related question is how much variability can be expected for a pair of samples as a function of time between the samples and how can that variation be explained. For iris, this question has not been carefully explored and to date, datasets suitable for such exploration have not been available. A key consideration is repeated instances of iris images densely packed in time that enable detailed exploration of noise; prior datasets, such as those from Notre Dame[3] , SMU[4] and Clarkson[5] have made collections several times per year on multiple subjects. In this paper we present an evolving dataset with collections conducted daily over periods of months. During the past year, for reasons beyond our control, we were only able to collect on a single subject. We hope to expand the subject pool in the future and are continuing data collection. We are publishing initial results from the work at this time and are open to recommendations on future collection/analysis ideas from the iris recognition community.

2. Description of Data and Analysis

IrisDaily-NIST is a dataset of iris images in which we tried to collect daily (we missed some days) from a single consented subject under a protocol approved by the NIST IRB.

The images were collected using a commercial, off-the-shelf, iris camera that captures left and right iris images near simultaneously. The capture software is the test harness associated with the IREX-VII framework; the IREX-VII wrapper for the camera is based upon the example code provided by the manufacturer and uses the camera control libraries provided by the manufacturer.

On each day of capture two pairs of left/right iris images were captured in quick succession in an office like environment, with less than a minute between the captures. The images are labeled with eye and a timestamp in the form <anonymousID>_YYYYMMDDHHmmSSsss_<L or R>.bmp, where YYYY is a four digit year, MM is a two digit month, DD is a two digit day of the month, HH is a two digit hour (24 hour clock), mm is a two digit minute, SS is a two digit second, and sss is a three digit milli-second. There were no intentional changes in the environment over the course of the collection:

- The camera and subject positions changed by less than 10 cm from nominal day to day.
- The room lights were not changed.
- The light diffusing window blinds were closed for all captures.
- The camera and subject positions were at the end of the room furthest from the window.
- There was a single subject who is well habituated with iris data collection in general as well as the specific camera.

There were also no special efforts taken to control the environment:

- Sunlight levels changed with time, date, season and weather.

Table 1 summarizes some important characteristics of the dataset. In the table FTE means *failure to enroll*; in this case the algorithm was able to process all of the images. It is important to note that the numbers in this table represent a snapshot of the dataset at a particular point in time. Additional data is being collected as this is written and will be added to the dataset.

Figure 1 is a typical image from the dataset.

Table 1. Description of IrisDaily-NIST Dataset

Number of iris images	396
Left Eye Images	198
Right Eye Images	198
Number FTE, iris2pi	0
Date Range	07/27/2020-12/02/2020
Time span (days)	128
Number days when images collected	99
Image type	iris
Image size	640x480
Illumination	near infrared (NIR)
Camera	IrisID-7111

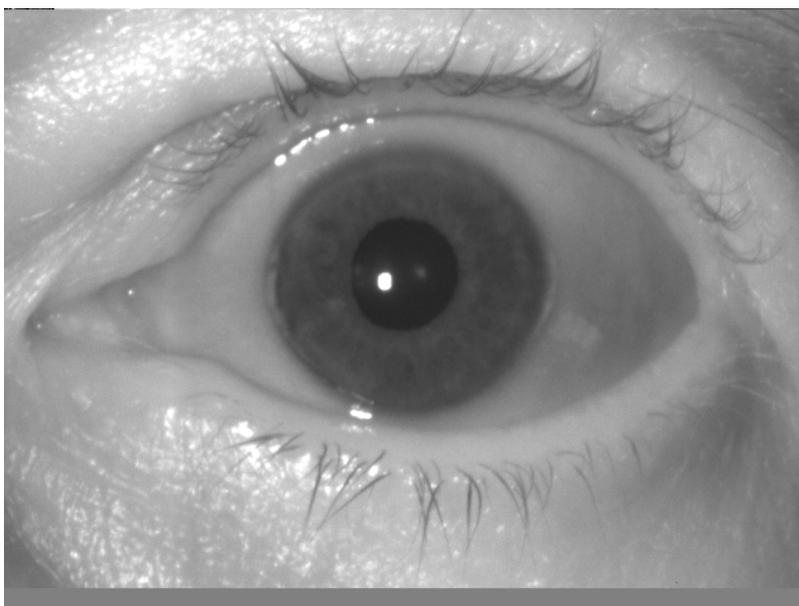


Fig. 1. Example of an image from the IrisDaily dataset. The subject is a NIST staff member who provided informed consent under a protocol approved by the NIST IRB; the subject also provided a release for the use of the image in publications.

The iris images were processed in pairs to generate comparison scores using a commercial version of the iris2pi iris recognition algorithm whose internals have been described in detail by Daugman[6] [7] [8] [9] and variants of which have been implemented by academics, see for example Masek[10].

We used R[11] to produce the plots and statistical tests presented here.

In the plots below, the comparison scores are raw scores as reported by the algorithm. The iris2pi scores are fractional Hamming distances which are dissimilarity scores. The fractional Hamming distances have not been normalized using the Daugman correction for number of bits compared (equation 6 in [8]). We are reporting the un-normalized scores to prevent variations in eyelid occlusion from contributing to comparison score variations through the Daugman correction.

3. Results

Figure 2 presents box-plots of the comparison scores for all comparisons between the images in IrisDaily-NIST for the iris2pi algorithm. As expected for high quality iris images compared using competent algorithms such as iris2pi[12], the separation between mated and non-mated scores is excellent. Figure 3 also presents the same data in histogram form. The perfect separation between mated and non-mated cases gives us assurance that ground-truth errors are absent.

Figure 4 shows the data from figure 3 binned by the time delta in days between the image acquisitions. We note that the only case where there is a significant difference in the box plots is for case of same day image pairs. This is consistent with observations in the literature, in particular the guidance in Appendix A of Mansfield and Wayman[13]: “Generally, 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.”

We can explore the differences in the distributions in more detail using qq-plots[14] as seen in figure 5. Here we note that ALL of the non-mated plots overlay as do all of the mated plots with the exception of the same day plot. This indicates that, with the exception of the mated, same day comparisons, all the mated plots have very similar distributions; the non-mated plots also have very similar distributions. Of course the non-mated and mated comparison scores come from distributions of different shape and location, as we know from the histograms of figure 3.

A more important observation is that even with a commercial grade camera deployed in an office setting with no intentional variation in ambient, with a single subject who is cooperative and well habituated, and with a short time interval between captures there is significant random variation in the mated comparison scores.

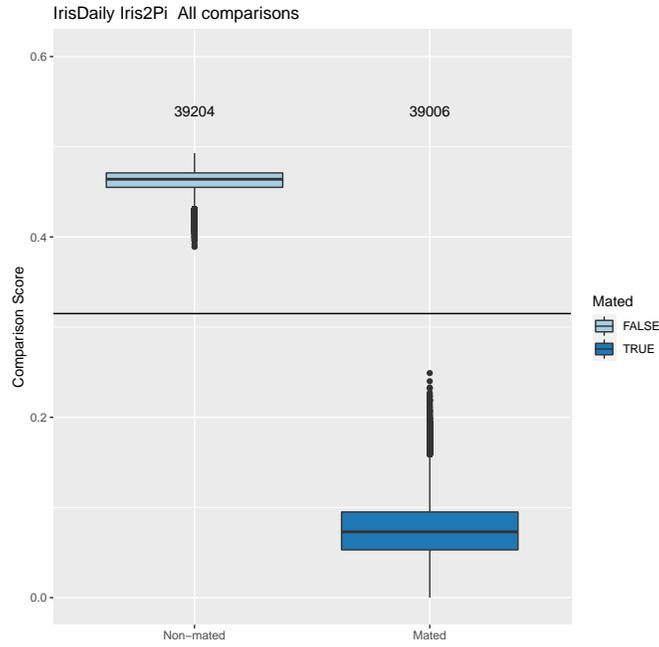


Fig. 2. Box-plots of comparison scores for mated and non-mate image pairs from the IrisDaily-NIST dataset. The numbers above the boxes indicate the number of comparison scores in the boxplot. This plot presents comparisons of all images to all images in the dataset. The horizontal line indicates a fractional Hamming distance, 0.315, that corresponds to a false match rate of approximately 1:1M on the basis of previously published results[8].

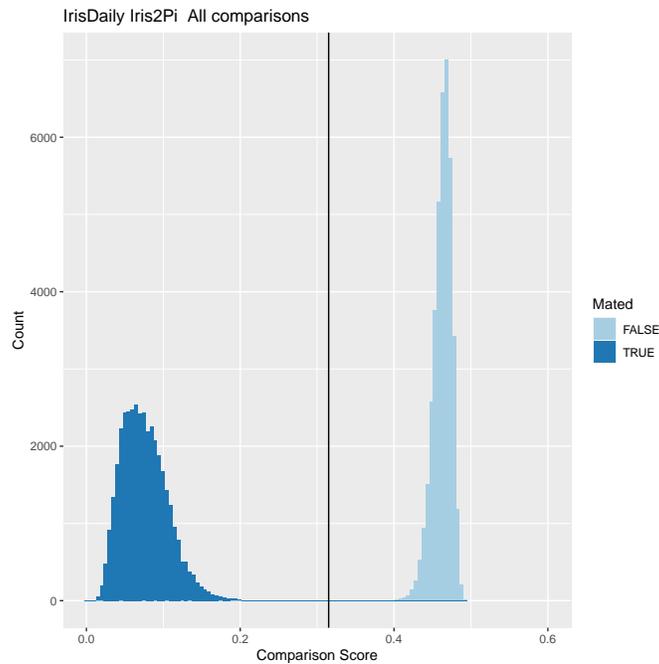


Fig. 3. Box-plots of comparison scores for mated and non-mate image pairs from the IrisDaily-NIST dataset. The numbers above the boxes indicate the number of comparison scores in the boxplot. This plot presents comparisons of all images to all images in the dataset. The vertical line indicates a fractional Hamming distance, 0.315, that corresponds to a false match rate of approximately 1:1M on the basis of previously published results[8].

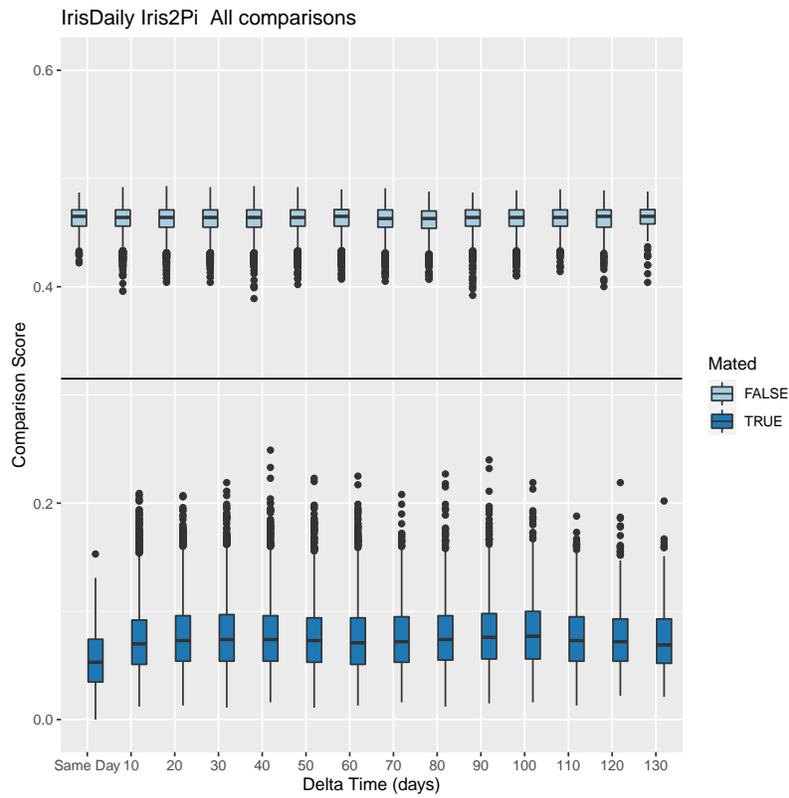


Fig. 4. Box-plots of iris2pi comparison scores by time delta. This plot presents comparisons of all images to all images in the dataset. The horizontal line indicates a fractional Hamming distance, 0.315, that corresponds to a false match rate of approximately 1:1M on the basis of previously published results[8].

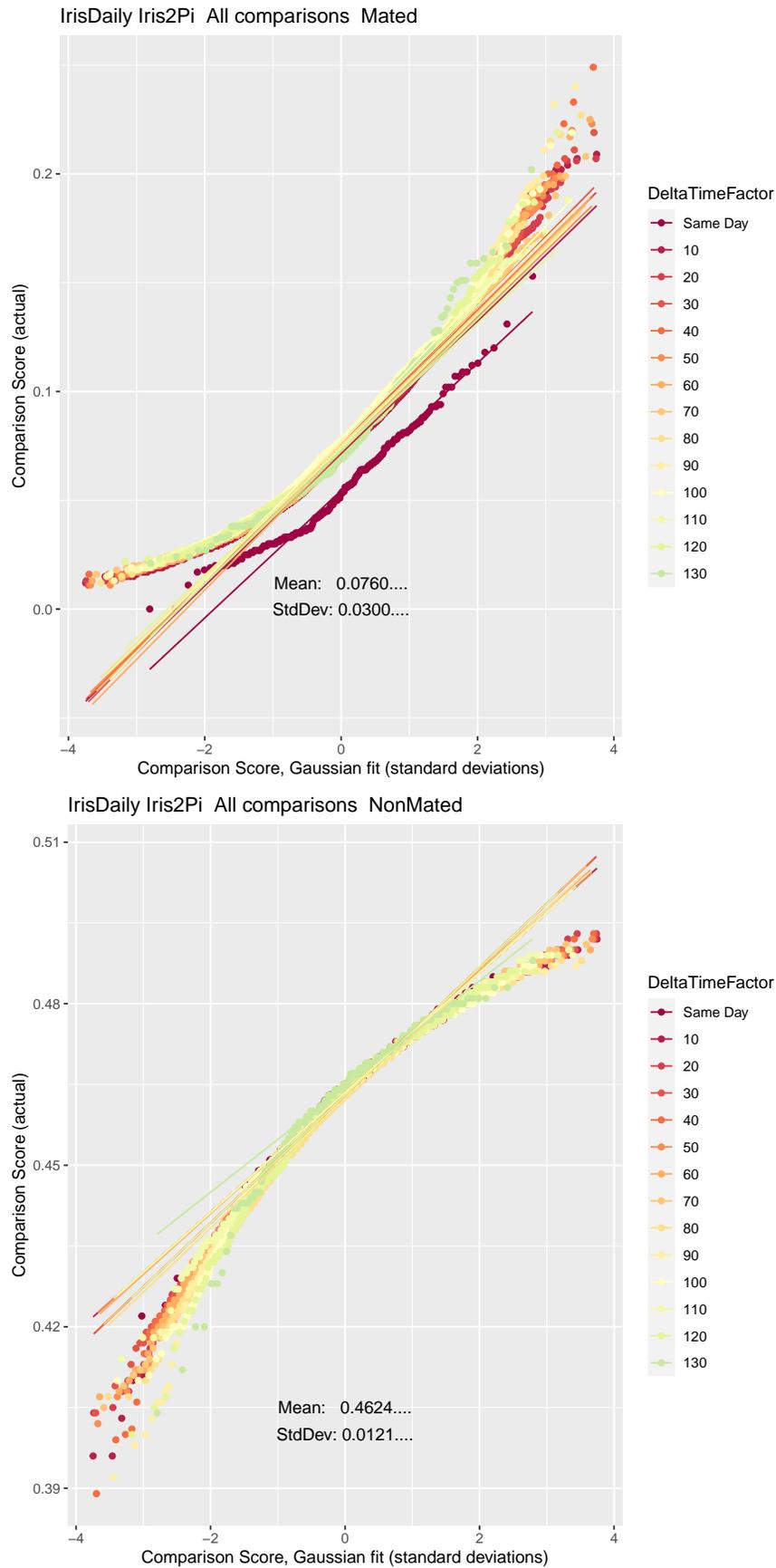


Fig. 5. The means and standard deviations are those for the data in aggregate. The lines represent fits to Gaussian distributions for each Delta Time factor. Deviations of the points from the lines indicate that the data is not normally distributed. Upward curvature indicates fat tails; downward curvature indicates lean tails. **Top:** QQ-plots of mated comparison scores. **Bottom:** QQ-plots of non-mated comparison scores.

4. Future Plans

- Continue to collect data almost every day and incorporate the new data into the dataset and future analyses.
- Collect data on a finer time scale to determine the minimum time interval between same day images required to make the same day mated qq-plot overlay the other mated qq-plots in figure 5.
- Analyze the data with additional algorithms by adding this dataset to our ongoing IREX evaluations.

Acknowledgments

Our thanks to Patrick Grother, Greg Fiumara, George W. Quinn, Bethany Retton and Craig Watson for helpful comments on the manuscript.

References

- [1] Grother PJ, Matey JR, Tabassi E, Quinn GW, Chumakov M (2013) IREX VI-Temporal stability of iris recognition accuracy. NIST, Technical report. URL doi.org/10.6028/NIST.IR.7948.
- [2] Matey JR, Matey JR, Quinn GW, Grother PJ (2021) Forensic Iris Review: 2020. NIST, Technical Report TN-xxxx, to be published.
- [3] Phillips PJ, et al. (2009) Frvt 2006 and ice 2006 large-scale experimental results. *IEEE transactions on pattern analysis and machine intelligence* 32(5):831–846.
- [4] Webb J, Etter DM, Barboza V, Sharp E (2016) Iris recognition using cross-spectral comparison. *2016 50th Asilomar Conference on Signals, Systems and Computers (IEEE)*, pp 434–438.
- [5] Das PD, Holsopple L, Rissacher D, Schuckers M, Schuckers S (2020) Iris recognition performance in children over time: A longitudinal study. *Submitted to The 2nd Workshop on Demographic Variation in the Performance of Biometric System, WACV 2020*, .
- [6] Daugman J (2001) Iris recognition. *American scientist* 89(4):326–333.
- [7] Daugman J (2004) How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on* 14(1):21–30. <https://doi.org/10.1109/TCSVT.2003.818350>
- [8] 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.
- [9] Daugman J (2016) Evolving methods in iris recognition, . URL http://www.cse.nd.edu/BTAS_07/John_Daugman_BTAS.pdf.
- [10] Masek L (2003) *Recognition of human iris patterns for biometric identification*. Ph.D. thesis. Master’s Thesis, Univesity of Western Australia, .
- [11] 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/>.
- [12] Daugman J (2002) How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology* 14(1):21–30.
- [13] Mansfield AJ, Wayman JL (2002) Best practices in testing and reporting performance of biometric devices, NPL Report CMSC 14/02 , . Technical report. URL https://www.kisa.or.kr/uploadfile/as-is/060405-BestPractices_v2_1-_-1153808236433.pdf.
- [14] Marden JI, et al. (2004) Positions and qq plots. *Statistical Science* 19(4):606–614.