NISTIR 8307

Interoperability Assessment 2019: Contactless-to-Contact Fingerprint Capture

John Libert John Grantham Bruce Bandini Kenneth Ko Shahram Orandi Craig Watson

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8307



NISTIR 8307

Interoperability Assessment 2019: Contactless-to-Contact Fingerprint Capture

John Libert Bruce Bandini Kenneth Ko Shahram Orandi Craig Watson Information Access Division Information Technology Laboratory

> John Grantham Systems Plus, Inc. Rockville, MD

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8307

May 2020



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

National Institute of Standards and Technology Walter Copan, NIST Director and Under Secretary of Commerce for Standards and Technology 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, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

National Institute of Standards and Technology Interagency or Internal Report 8307 Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8307, 48 pages (May 2020)

> This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8307

VERSION HISTORY

Date	Activity

Disclaimer

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

The National Institute of Standards and Technology Research Protections Office (RPO) reviewed the protocol for this project and determined it meets the criteria for exempt human subjects research as defined in Department of Commerce Regulations, 15 CFR 27, also known as the Common Rule for the Protection of Human Subjects (45 CFR 46, Subpart A).

Acknowledgement

We would like to acknowledge and thank the Federal Bureau of Investigation's (FBI) Criminal Justice Information Services (CJIS) Division as well as other key partners at the FBI who provided support to National Institute of Standards and Technology (NIST) in this contactless fingerprint study.

ABSTRACT

This report describes computational comparison measurements as well as match error analysis applied to fingerprint images collected by researchers of the National Institute of Standards and Technology (NIST) in May of 2019. Fingerprint images from 200 Federal Employees were collected using both contact and contactless acquisition devices, including both stationary devices and mobile smartphone devices. Match scores from two different matchers were used to calculate False Negative Identification Rate (FNIR) and False Positive Identification Rate (FPIR) at two operational thresholds appropriate to each of the matchers. These values are displayed in graphical form across eight devices for each of ten finger combinations. Overall error performance is compared via Area Under Curve across devices for each of ten finger combinations as well as across finger combinations separately for each of the eight devices. Contact-to-contact interoperability remains superior to contactless-to-contact, but multiple finger combinations improve match performance of contactless devices.

KEYWORDS

contactless fingerprinting; fingerprint interoperability; minutiae correlation; image structural similarity; fingerprint image registration

TABLE OF CONTENTS

1	B	ackground	1
2	Р	urpose	2
3	D	ata	3
	3.1	Collection Procedure	3
4	R	esults	4
	4.1	Feature Measurement	4
	4.2	Matcher Tests: Error Analysis by Threshold	11
	4.3	Match Error Performance Summaries	24
	4.4	Matcher Tests: Area Under the ROC Curve	25
5	D	viscussion and Conclusions	37
	5.1	Metrics	37
	5.2	Matching	37
	5.3	Other Conclusions and Observations	37
6	R	eferences	38

LIST OF TABLES

Table 1 – Abbreviations	v
Table 2 – TPIR/FPIR at threshold 8 500 (%)	24
Table 3 – TPIR/FPIR at threshold 4 100 (%)	24

LIST OF FIGURES

Figure 1 – Dev01 and Dev02 are Appendix F certified contact devices, Dev03 and Dev04 are stationary contactless devices, and Dev05 – 08 are smartphone applications. (Red dashed lines demark PIV limits; Green dashed lines demark Appendix F limits.)
Figure 2 – Distributions of number of corresponding minutia pairs between the listed device and Dev01. More corresponding minutiae pairs relate to better matches between pairs of images. For Dev01, a comparison is made between two impressions taken from the same device
Figure 3 – Mean displacement of corresponding minutiae pairs
Figure 4 – Normalized overlap between probe and exemplar fingerprint images
Figure 5 – SSIM shows the structural similarity between fingerprint impressions captured with the same device (Dev01) and different capture technologies (Dev02 – Dev08). Higher SSIM values indicate greater similarity
Figure 6 – Ridge orientation map correlation (Note: Range of correlation coefficient is from -1 to +1)
Figure 7 – Similarity score distributions as output by the minutiae correspondence tool. Scores below 500 overlap into the range of impostor scores generated by this modified matcher
Figure 8 – ANSI/NIST-ITL 1-2011 finger designation by number [4]. (Note that, in practice and in this document, finger designations typically include a leading zero so that all designations are two digits in length.)
Figure 9 – Illustration of decision errors relative to a selection threshold (<i>t</i>) applied to genuine and impostor score distributions (from [8])
Figure 10 – FNIR with 95 % confidence intervals plotted for two potential operating thresholds, using both matchers (Note that the Dev01 probes are from encounter #2 (E02) acquired on this device, while exemplars are from encounter #1 (E01) with the same device)
Figure 11 – Distributions of match scores by finger from the LE matcher, illustrating that highest scores are generally obtained from index and middle fingers (02, 03, 07, 08). This is the case for both contact and contactless devices14
Figure 12 – Distributions of match scores by finger from the COM matcher, illustrating that highest scores are generally obtained from index and middle fingers (02, 03, 07, 08). This is the case for both contact and contactless devices15
Figure 13 – Finger 02 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers
Figure 14 – Finger 07 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers
Figure 15 – Fingers 02 and 03 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers
Figure 16 – Fingers 07 and 08 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers
Figure 17 – Four fingers of right hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using LE matcher, shown beside the first three fingers of the right hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using COM matcher
Figure 18 – Four fingers of left hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using LE matcher, shown beside the first three fingers of the left hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using COM matcher

Figure 19 – Index fingers of each hand matched to encounter #1 (E01) samples acquired on the control device, using both matchers	Dev01,
Figure 20 – Two + Two finger combination appears to produce the best performance for several devices, though wit	h some
disparity between the two matchers.	
Figure 21 – Eight finger combination generates lower error for most devices using the LE matcher, shown beside i	the six-
finger combination, which generates lower error for most devices using COM matcher	22
Figure 22 – FNIR with Dev03 prints as exemplars probed by prints from other devices, using LE matcher only	23
Figure 23 – AUC for single finger from both matchers	26
Figure 24 – AUC for index finger from right hand using both matchers	27
Figure 25 – AUC for index finger from left hand using both matchers	27
Figure 26 – AUC for index and middle fingers from right hand using both matchers	28
Figure 27 – AUC for index and middle fingers from left hand using both matchers	28
Figure 28 – AUC for four fingers from right hand using LE matcher shown beside AUC for first three fingers usin	ig COM
matcher	29
Figure 29 – AUC for four fingers from left hand using LE matcher shown beside AUC for first three fingers from left	ft hand
using COM matcher	29
Figure 30 – AUC for index fingers from both hands using both matchers	30
Figure 31 – AUC for index and middle fingers from both hands using both matchers	30
Figure 32 – AUC for all eight fingers from both hands using LE matcher shown beside AUC for the first three finger	rs from
both hands using COM matcher	31
Figure 33 – AUC for all finger combinations with Dev01 using both matchers	32
Figure 34 – AUC for all finger combinations with Dev02 using both matchers	33
Figure 35 – AUC for all finger combinations with Dev03 using both matchers	34
Figure 36 – AUC for all finger combinations with Dev04 using both matchers	34
Figure 37 – AUC for all finger combinations with Dev05 using both matchers	35
Figure 37 – AUC for all finger combinations with Dev05 using both matchers Figure 38 – AUC for all finger combinations with Dev06 using both matchers	35
Figure 37 – AUC for all finger combinations with Dev05 using both matchers Figure 38 – AUC for all finger combinations with Dev06 using both matchers Figure 39 – AUC for all finger combinations with Dev07 using both matchers	35 35 36

TERMS AND DEFINITIONS

Table 1 – Abbreviations

ANSI	American National Standards Institute
AUC	Area Under (ROC) Curve
DET	Detection Error Tradeoff
DUT(s)	Device(s) under test
FBI	Federal Bureau of Investigation
FNIR	False Negative Identification Rate
FPIR	False Positive Identification Rate
FTIR	Frustrated Total Internal Reflection – optical principal employed by some fingerprint acquisition devices
NIST	National Institute of Standards and Technology
ррі	Pixels per inch (the customary unit of sampling for digital fingerprints)
ppmm	Pixels per millimeter
ROC	Receiver Operating Characteristic
TPIR	True Positive Identification Rate

Executive Summary

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8307

Contactless acquisition of fingerprints presents a fundamental departure from legacy, contact-based capture technologies. Despite various forms and degrees of distortion possible with contact capture, an outstanding characteristic of legacy fingerprint acquisition is that ridges are unambiguously recorded by the contact process itself. The representation of the friction ridge surface might suffer certain defects due to insufficient pressure, too much pressure, over- or under-inking, sliding, etc., but the position where the ridge appears on the paper or in the digital image will be close to the truth. For those contactless devices examined by NIST, all start with an optical representation of an illuminated finger surface – a photographic image, itself a projection of three-dimensional (3D) structure onto a two-dimensional (2D) plane. With contact capture, the lights and darks of the images bear a direct relationship to the topographic lows and highs of the friction ridge surface, thus yielding a first-order representation. With contactless capture, lights and darks are the result of the 3D surface being modeled by illumination as reflection and shadow, thus lacking the tight coupling between image and surface features characteristic of contact capture. Furthermore, with contactless capture unspecified image processing procedures are applied to the source photographic representation to create a grayscale or binary image that more closely resembles that of a legacy fingerprint impression, thus yielding a third-order representation.

For this investigation, 200 Federal employee volunteers were recruited to participate in a data collection to have their fingerprints acquired with both contact and contactless devices. We note that the contactless devices used in the study are current models (as of May 2019), thus a fair representation of the state-of-the-art for contactless capture at the time of the investigation. Contact devices included representatives of two different image-forming technologies to provide an interoperability baseline. Contactless devices included two stationary models and four software applications running on several models of smartphone.

Following procedures detailed in NIST Special Publication 500-305 [1], images from all devices included in the collection are compared to corresponding exemplars from a legacy contact device which served as a control. Selected computational comparison measurements are applied to image pairs, notably measurement of scale (sample rate), displacement of corresponding minutiae, structural similarity, and matcher-based similarity. Considering these metrics, we find contact-to-contact performance to be best, followed by stationary contactless, with mobile contactless devices performing less capably. Notably, control of scale (or sample rate) remains a challenge for mobile contactless devices.

Early NIST matcher testing and device interoperability evaluation among contactless and contact devices was limited to small data sets from several disparate fingerprint data collections [1]. While still small, this investigation includes a large gamut of devices in one study, sampling the same set of fingerprint donors for all eight devices under test (DUTs). The main purpose of this investigation is to further test the hypothesis that multiple finger matching can significantly improve the error performance of contactless fingerprint matching.

To this end we use two different matchers to evaluate False Negative Identification and False Positive Identification Rates (FNIR and FPIR) in probing a database of contact exemplar mates with a background of three million known non-mate contact impressions. We evaluate error at two operating thresholds appropriate to each of the matchers and compare FNIR across the eight devices for each of ten finger combinations. Also, we evaluate error via computation of Area Under the ROC Curve (AUC) both across devices for each finger combination and across finger combinations for each of the eight DUTs. With both analyses, we see the lowest error for contact-to-contact comparisons, with FNIR mostly below 1.0 % and with one stationary contactless device approaching this level of performance. For the most part, mobile contactless devices show greater error, but several mobile devices begin to reach error rates of 5 - 10 % with some finger combinations. We confirm our hypothesis that multiple finger matching significantly improves match error performance of contactless devices.

1 Background

Contactless acquisition of fingerprints presents a fundamental departure from legacy, contact-based capture technologies. The friction ridge surface that comprises the fingerprint is a three-dimensional (3D) topography that is present on the three-dimensional curved surface of a finger. The legacy-compatible fingerprint impressions output from contactless devices are two-dimensional (2D) representations of this three-dimensional structure. Currently, there are no accepted industry standards for 3D representations of fingerprints that demonstrate compatibility with legacy fingerprint databases.

For contact acquisition, the third dimension of the fingerprint ridge structure is effectively removed through contact with the surface of the recording medium, be that paper, the platen of an Frustrated Total Internal Reflection (FTIR) optical device, or other flat surface used in capacitive, electro-optical, and ultrasonic devices. Even though the finger surface is curved, it is flattened by deformation of the finger through pressure against the capture surface. The friction ridge structure is acknowledged to the extent that topographic high points (i.e., ridges) are recorded and lows (i.e., furrows/valleys) are not. For a rolled acquisition, the three-dimensional structure of the finger is unwrapped onto a 2D plane by rolling the finger across the capture surface. For a plain (flat) capture, the finger pad is flattened against the capture surface. Despite various forms and degrees of distortion possible with contact capture, a characteristic of legacy fingerprint acquisition is that ridges are unambiguously recorded by the contact process itself. The representation of the friction ridge surface might suffer certain defects due to insufficient pressure, too much pressure, over- or under-inking, sliding, etc., but the position where the ridge appears on the paper or in the digital image will be close to the truth.

In inked fingerprint capture, the print on the card is a representation of the friction ridge surface. When the card is scanned, the digital image is a representation of the inked impression – a second-order representation of the ridge structure of the finger surface. For FTIR and other "direct capture" devices, the digital image is a first-order representation of the friction ridge surface. In this regard, it is noted that, for ink-on-card fingerprints, most defects that affect quality occur before the digitization process (i.e., at the point that the inked impression is recorded on the paper). Additional error could result from conversion of the inked impression to digital form via optical scanning, though for inked impressions, the quality is mainly determined before optical scanning is brought to bear.

For those contactless devices examined by NIST, all start with an optical representation of an illuminated finger surface – a photographic image of some sort, itself a projection of the 3D structure onto a 2D plane. With contact capture, the lights and darks of the images bear a direct relationship to the topographic lows and highs of the friction ridge surface, thus yielding a first-order representation. With contactless capture, lights and darks are the result of the 3D surface being modeled by illumination as reflection and shadow, thus lacking the tight coupling between image and surface features characteristic of contact capture. Furthermore, with contactless capture unspecified image processing procedures are applied to the source photographic representation to create a grayscale or binary image that more closely resembles that of a legacy fingerprint impression, thus yielding a third-order representation. As such, one objective of the present investigation is to evaluate the interoperability of these contactless representations with legacy fingerprint impressions in the context of current matchers.

2 Purpose

Early NIST matcher testing and device interoperability evaluation among contactless and contact devices was limited to small data sets from several disparate fingerprint data collections [1]. In the present instance, the dataset has been expanded significantly but remains small relative to the massive legacy datasets that have been established over the decades. This investigation also provides a larger gamut of devices in one comparative study, utilizing the same set of subjects as fingerprint donors. Previous tests suggested that multiple finger combinations might increase the performance level of contactless towards potential operational utility. As most contactless devices capture four fingers at a time, multiple finger matching is more practical than single finger capture and matching. A pair of four-finger captures enables comparison of any multiple finger combination exclusive of thumbs¹. Thus, the main purpose of the present investigation is to further test the hypothesis that multiple finger matching can substantially improve the accuracy of contactless fingerprint matching.

¹ For many contactless devices, thumb capture has been ergonomically cumbersome in our experience.

3 Data

3.1 Collection Procedure

We recruited 200 Federal employee volunteers from among the staff at NIST. Each participant was assigned a numerical identifier for anonymous coding of fingerprints and other data collected, as described below. A Quick Response (QR) code and numeric representation of the participant identifier was printed on a sheet that the participant carried from station to station.

Upon reporting at a prearranged time, each participant was handed a project description (informed consent document), asked to review it, and to agree to contribute their fingerprint biometric samples to research being conducted to evaluate interoperability among various fingerprinting devices. No signature was required, as data collected is to remain anonymous (i.e., not traceable to individual contributors).

Participants (only a few at a time in the room) circulated among data collection stations as follows:

- 1- Contact plain print collection (four-finger slaps);
 - a. $FTIR^2 Dev01^3$ (two encounters)
 - i. Encounter #1 is used as the exemplar and included in the gallery for all subsequent comparisons;
 - ii. Encounter #2 is used as the control probe for the baseline contact device (Dev01).
 - b. Electro-luminescent Dev02
- 2- Stationary (desktop) contactless devices (two);a. Dev03, Dev04.
- 3- Smartphone contactless applications (four);
 - a. Dev05 08.
- 4- Hand color measurement;
 - a. Digital Single-Lens Reflex (DSLR) camera photographing the palmar side of either hand next to a calibrated commercial color checking device.

Fingerprints captured as four-finger "slap" impressions were segmented into individual fingerprint images, where applicable, by the device under test (DUT). Contact slaps of the control device, Dev01 were segmented via a software tool and verified visually for accuracy.

It should be noted that the contactless devices used in the study are current models⁴, thus good indication of the state-of-the-art for contactless capture at the time of the collection.

² FTIR = Frustrated Total Internal Reflection – an optical property utilized by some fingerprint capture devices to differentiate points of contact from those of non-contact (or reduced contact) in representing the friction ridge relief of a fingerprint.

³ Device models are anonymized in this report and are assigned the generic identifiers Dev01 – Dev08.

⁴ Current as of May 2019, though several may have undergone some modification as of publication of this report.

4 Results

4.1 Feature Measurement

Following procedures detailed in NIST Special Publication 500-305 [1], images from all devices were compared to corresponding exemplars from Dev01, encounter #1. Selected measurements are shown in the boxplot figures below. Similarity scores are included in the feature set, but additional metrics included below reveal more specific information about image quality relative to contact exemplars. We include several of the most important metrics that will likely be included in a proposed testing protocol for device performance evaluation for potential certification. Other measurements described in [1] could potentially be included in a forthcoming proposed testing protocol and performance specifications.

4.1.1 Sample Rate

Sample rate continues to be a challenge for contactless devices, as can be seen in Figure 1. As a byproduct of the registration of probe image to contact exemplar, we acquire a measure of the scale difference between the two images under comparison. The resulting scale factor, which varies around 1.0, is converted to a sample rate in pixels per pinch⁵ (ppi) by multiplying by 500 ppi. Distributions of sample rate measurements are shown in Figure 1. The distributions of Dev01 (encounter #1) and Dev02 fit within the 500 ppi \pm 5 ppi required by the Appendix F [2] specification, and Dev03 largely fits within the 500 ppi \pm 10 ppi requirement for PIV [3]. All but one device has a median sample rate of 500 ppi, but most of the contactless devices (Dev04-08) show portions of their distributions exceeding the PIV limits by varying degrees⁶.



Figure 1 – Dev01 and Dev02 are Appendix F certified contact devices, Dev03 and Dev04 are stationary contactless devices, and Dev05 – 08 are smartphone applications. (Red dashed lines demark PIV limits; Green dashed lines demark Appendix F limits.)

⁵ Resolution values for friction ridge imagery are specified in pixels per inch (ppi) throughout this document. This is based on widely used specification guidelines for such imagery and is accepted as common nomenclature within the industry. SI units for these will be presented only once. 500 ppi is approximately 19.7 ppmm.

⁶ Boxplots indicate median values by a horizontal line; notches indicate 95 % confidence limits about the median; the box contains 50 % of the cases while the vertical lines, "whiskers", each denote 24.65 % of the cases; extreme outliers are omitted from these plots.

4.1.2 Number of Minutiae Pairs

One method of establishing the degree of fidelity between two fingerprint samples is demonstrating detection of corresponding minutiae pairs between those two fingerprint samples. For this, a commercial-grade minutia detector was employed. This matcher-based tool was modified to enable the output of lists of *x*, *y* coordinates classified by the modified matcher as "corresponding". One aspect of fingerprint quality is indicated by the number and proximal agreement of such pairs. Figure 2 exhibits the distributions of numbers of corresponding minutiae pairs found between probe images and the contact exemplars of the first encounter using the contact device Dev01.

In Figure 2, we see that the largest number of corresponding minutiae pairs are found in comparing the two encounters acquired with Dev01 – the FTIR contact device. With both exemplar and probe acquired using the same device, this might well be expected, as we observe a lower median for a second contact device (Dev02) that employs a different contact-based capture technology. Dev03 and Dev04, stationary contactless devices, show further depression in median numbers of corresponding minutiae pairs when compared to mated contact exemplars from Dev01. Mobile (smartphone) contactless devices 05 - 08 show comparably lower numbers of minutiae pairs corresponding those of contact exemplars.



Figure 2 – Distributions of number of corresponding minutia pairs between the listed device and Dev01. More corresponding minutiae pairs relate to better matches between pairs of images. For Dev01, a comparison is made between two impressions taken from the same device.

4.1.3 Mean Minutiae Displacement

In addition to the number of corresponding minutiae pairs, it also is useful to examine the measured spatial displacement (distance) between the corresponding minutiae as computed during the registration process. We see the distributions of these measurements in Figure 3. The ordering in this metric is similar (though inverse) to that of the previous metric. The smallest average offset between corresponding minutiae is found with exemplars and probes from the same contact device, Dev01, followed by the second contact device, Dev02. Stationary contactless devices, Dev03 and Dev04, are next in line, followed by the mobile contactless devices, Dev05 - Dev08, which exhibited higher medians for this metric as well as larger variability in distributions (i.e., larger boxes and longer "whiskers").



Figure 3 – Mean displacement of corresponding minutiae pairs

4.1.4 Normalized Overlap

As many image comparisons, incuding matching, can be performed only to the extent that friction ridge regions of the images under comparison overlap, it is useful to measure the areas of overlap between probe and exemplar fingerprints as described in [1]. In Figure 4 we see medians over 90 % for the two contact devices and the two stationary contactless devices, as well as low variability relative to that of the mobile contactless devices. Dev07 and Dev08 have medians around 90 % as well, however, with greater variability. Dev03 attempts to capture a rolled equivalent image, hence generates fingerprints that are generally of much larger area than contact plain impressions used as exemplars. Mobile devices often have reduced area of overlap with contact plain impressions, which we suspect is due to geometric misalignment of the fingerprint core. That is, we observe that while the contactless fingerprint may exhibit good ridge definition, the print is "rolled" off center relative to the plain impression with which it is being compared, hence yielding only a small area of overlap for comparison.



Figure 4 – Normalized overlap between probe and exemplar fingerprint images

4.1.5 Structural Similarity Index (SSIM)

The computation of the SSIM is described in [1]. The method scores each of the corresponding blocks of the pair of images under comparison. The computation is directed toward evaluating the oriented contrast structures of the corresponding blocks. The similarity measure of all blocks is then pooled to yield a single figure of merit, the SSIM, quantifying the structural similarity of the pair of images. This metric is applied only to registered regions of overlap.

In Figure 5, SSIM shows the greatest structural similarity between fingerprint impressions captured with the same device (Dev01) and somewhat lesser similarity for comparison of contact impressions from the device employing a different capture technology (Dev02). SSIM reduces further for contact-to-contactless comparison, with lower values for stationary contactless (Dev03 and Dev04) and lowest SSIM for mobile contactless.



Figure 5 – SSIM shows the structural similarity between fingerprint impressions captured with the same device (Dev01) and different capture technologies (Dev02 – Dev08). Higher SSIM values indicate greater similarity.

4.1.6 Correlation of Ridge Orientation Maps

Once images are registered, overlapping regions cropped from source images may be compared with respect to similarity of ridge flow at corresponding points in the images under comparison. As described in [1], the local orientation of a ridge at each pixel of an image can be calculated to produce an orientation map of ridge direction in radians. The orientation maps for the two fingerprints under comparison are compared by 2D correlation.



Figure 6 – Ridge orientation map correlation (Note: Range of correlation coefficient is from -1 to +1)

Figure 6 shows the distributions of this metric for probe images of each device compared to contact exemplars, encounter #1 of Dev01. Contact devices Dev01 and Dev02 exhibit both the highest median scores and the lowest variability. Next in line are the two stationary contactless devices, Dev03 and Dev04. Smartphone devices have medians in the 0.5 - 0.65 range⁷ and higher variability.

⁷ Dev06 exhibits spuriously high metric performance as many prints could not be acquired by this device leaving only the "best quality" for measurement. Match scores presented later applied a correction for Failures to Acquire, (FTAs).

4.1.7 Similarity Score

The correlation patterns exhibited in the previous tests relate well to similarity scores output by the matcher modified to identify corresponding minutiae⁸. Distributions are shown in Figure 7. Note that this modified matcher was used to select corresponding minutiae between two images under comparison. In addition to lists of paired minutiae used as control points for registration, the modified matcher also yields a similarity score. Unlike matcher experiments described in Sections 4.2, 4.3, and 4.4, only a small number of impostor matches were evaluated using this modified matcher. The median similarity score for a distribution of 5 096 known non-mate comparisons was around 200, with the bulk of the distribution falling below 500. Some of the devices considered here have part of their distributions overlapping this impostor score range, but most have the bulk of the cases above the 500 mark. Notably, the trend in similarity scores among the different devices is consistent with performance of the other metrics evaluated in the preceding sections.



Figure 7 – Similarity score distributions as output by the minutiae correspondence tool. Scores below 500 overlap into the range of impostor scores generated by this modified matcher.

⁸ Note that match scores summarized here are output from the minutiae correspondence tool and are different from those used in the error analysis described later in this report.

4.2 Matcher Tests: Error Analysis by Threshold

4.2.1 Finger Numbering

For readers unfamiliar with the standard method of finger designation [4], we supply Figure 8 showing the unique numbering for fingers on right and left hands⁹. In the present study, we exclude thumbs (i.e., fingers 01 and 06).



Figure 8 – ANSI/NIST-ITL 1-2011 finger designation by number⁹ [4].

4.2.2 Matchers

While it would be impractical to perform experiments involving large numbers of human subjects for device certification testing, such data will prove essential for calibrating device responses to the test artifacts under development as part of our ongoing work on contactless device testing. Toward this end, we used a state-of-the-art matcher, configured to mirror operational scenarios in law enforcement applications, to examine match error rates for each of the devices against exemplars sampled on an FTIR contact device. To this matcher we have assigned the label Law Enforcement (LE) matcher. It is considered a "10-print matcher" capable of scoring any combination of fingers in a probe to corresponding fingers in a reference database and yielding a single "fused" score.

For comparison, we selected one of the PIV-compliant matchers used in the NIST MINEX [5] program. We term this single-finger matcher the Commercial or COM matcher. As this matcher yielded only a score for each single-finger comparison, it was necessary to fuse scores for multiple finger combinations. The simplest method was to simply average the scores. Due to the fact that the little fingers (fingers 05 and 10) scored substantially lower than the other three fingers in a four-finger slap (see Figure 11), we excluded them from multiple finger combinations.

⁹ Note that, in practice and in this document, finger designations typically include a leading zero so that all designations are two digits in length.



Figure 9 – Illustration of decision errors relative to a selection threshold (t) applied to genuine and impostor score distributions (from [8]).

Each matching experiment yields two distributions of scores (see Figure 9). For each device, we have a distribution of *N* genuine match scores of probe fingerprint images compared to their mated exemplars, and a distribution of *M* impostor, or non-mate, scores representing the highest score for each probe compared to 3 million known non-mates. In our scoring, N = M, as for each probe we have both a mate score and the highest impostor (known non-mate) score. At each selected score decision threshold, *t*, the False Negative Identification Rate (FNIR) is determined as the proportion of the genuine scores < *t*, and the False Positive Identification Rate (FPIR) is the proportion of impostor scores $\geq t$.

We examine FNIR and FPIR of various finger combinations of probe prints against an exemplar gallery of encounter #1 (E01) of Dev01 and a background of three million known non-mate records. For each device, we have approximately 1 590 images of each probe set¹⁰ submitted to matchers grouped in various finger combinations. Multiple-finger matches reduce the single-finger sample size to approximately 198 samples.

Rather than displaying error measurements using Receiver Operating Characteristic (ROC) or Detection Error Tradeoff (DET) curves, we chose to compare device performance by computing error at two operating thresholds. For the LE matcher, values appropriate for application of this matcher to searching large databases were set at 8 500 and 4 100. With our small dataset, the value 4 100 appeared to be a critical point where lower values resulted in increases in the FPIR, while the higher value of 8 500 was included for comparison. FPIR remained less than or equal to the upper 95 % confidence limit computed via the Rule of Three [6][7] (see Table 2 and Table 3), where the bounds of this interval are [0, 3/N].

The COM matcher yields scores in the range 0.0 - 1.0. Hence, to set thresholds to aid comparison between the two matchers, values were selected in this range that were calculated as proportional to those selected for the LE matcher. The value 0.40 was tested empirically as the lower limit beyond which we observed an increase in FPIR levels similar to that observed for the LE matcher. Hence, we set the two thresholds for the COM matcher at 0.60 and 0.40 to reflect those set for the LE matcher

In the plots of the following sections, we include FNIR values at both upper and lower thresholds but note that the best measure of performance is that at the threshold 4 100 for the LE matcher and 0.40 for the COM matcher. As the FPIR remained less than or equal to the 3/N limit prescribed by the Rule of Three [6][7] for all thresholds for each of the matchers, we display only the FNIRs.

¹⁰ Sample sizes varied slightly due to varying numbers of Failures to Acquire (FTAs) across devices.

4.2.4 Single Finger

Each capture consisted of eight fingers, four from each hand. From experience, we found that thumbs are not conveniently captured by some of the contactless devices examined, and most contactless devices are best suited to four-finger captures. Single-finger FNIRs are shown in Figure 10.



Figure 10 – FNIR with 95 % confidence intervals plotted for two potential operating thresholds, using both matchers (Note that the Dev01 probes are from encounter #2 (E02) acquired on this device, while exemplars are from encounter #1 (E01) with the same device).

Confidence intervals are determined by a bootstrap method in which the plotted value is computed by the first analysis of genuine and impostor score distributions and the error analysis repeated for each of 2 000 [9] replicate samples taken randomly with replacement from the input distributions of scores. The 95 % confidence limits demark the 0.025th and 98.025th quantiles of the distribution of FNIR replicate values.

We see that single finger results are likely spuriously low for most of the devices. This should not be surprising, as is seen graphically in Figure 11 and Figure 12 where it is evident that not all fingers generate high match scores. Single finger results are commonly presented to increase sample size for a test involving a small number of fingerprint donors. However, one of the principal motivations for the data collection described here is to examine match results for multiple finger combinations, even at the expense of lower statistical confidence. While the sample size is reduced substantially for multiple finger matching experiments increasing uncertainty, consideration of multi-finger matching is realistic from an operational point of view. All the devices currently under test by NIST capture four fingers at a time in their proposed application. Hence, it is reasonable to assess their utility in that context.



Figure 11 – Distributions of match scores by finger from the LE matcher, illustrating that highest scores are generally obtained from index and middle fingers (02, 03, 07, 08). This is the case for both contact and contactless devices.



Figure 12 - Distributions of match scores by finger from the COM matcher, illustrating that highest scores are generally obtained from index and middle fingers (02, 03, 07, 08). This is the case for both contact and contactless devices.

4.2.5 Index Finger Only

The single finger error results summarized in Section 4.2.4 could be misinterpreted in that it is unlikely that a single-finger capture would be applied with equal likelihood to any of the ten (or eight) fingers available. The most likely choice for a single finger capture would be the index finger. At least one contactless (smartphone) device being used in a study being conducted in the United Kingdom (UK) captures only a single index finger, as does a contact device currently used by police in the UK¹¹. Hence, we evaluate FNIR in Figure 13 for Finger 02 (index finger of the right hand) for both LE and COM matchers, and in Figure 14 for Finger 07 (index finger of the left hand).



Figure 13 – Finger 02 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers



Figure 14 – Finger 07 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers

¹¹ Pers. Comm., Richard Case, Forensics and Identity Team, Policing and Security Group, Counter Terrorism and Security Division Defense, Science and Technology Laboratory, Ministry of Defense, United Kingdom.

Interestingly, right- and left-hand performance results are not the same. Capture procedures are identical for the two hands, but there may be an order difference with typical capture of the right hand first. Unfortunately, most of the applications are fixed in such way that randomizing the order of capture is not possible without changes to the application software, which could only be performed by the manufacturer of a device.

4.2.6 Two Fingers

With the addition of the middle finger (03) to the index finger (02) of the right hand, the FNIR performance improves for all devices as see in Figure 15.



Figure 15 – Fingers 02 and 03 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers

Similar results were obtained for the index (07) and middle (08) finger of the left hand in Figure 16.



Figure 16 – Fingers 07 and 08 matched to encounter #1 (E01) samples acquired on the control device, Dev01, using both matchers

4.2.7 Three/Four Fingers

Here, the LE and COM matchers are treated somewhat differently. The LE matcher has its own internal fusion algorithm such that adding low scoring little fingers (05 or 10) does not degrade scores derived from the other three fingers. For the COM matcher, the scores of the three fingers are simply averaged, and the little fingers are dropped from the scoring.

In any case, we see that adding additional fingers reduces FNIR even more than that observed for two-finger combinations (02,03 or 07,08 as shown in section 4.2.6.) for some devices, as seen in Figure 17.



Figure 17 – Four fingers of right hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using LE matcher, shown beside the first three fingers of the right hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using COM matcher

As with the right hand, we see similar improvement when adding additional fingers to the index and middle fingers (see Figure 18).



Figure 18 – Four fingers of left hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using LE matcher, shown beside the first three fingers of the left hand matched to encounter #1 (E01) samples acquired on the control device, Dev01, using COM matcher

4.2.8 One + One Index Finger Combination

A potential two-finger combination might include the two index fingers. The results for this test case are shown in Figure 19. We see that both contact devices exhibit near zero error with both matchers. The stationary contactless device, Dev03, shows an FNIR of around 0.05 for both matchers. With the LE matcher Dev08, a smartphone application, performs fourth best with an FNIR just under 0.1 for the LE matcher.





4.2.9 Two + Two Finger Combination

Improved performance is observed for most devices using the index and middle fingers of both the right and left hand (i.e., fingers 02, 03 and 07, 08). FNIR approaches zero for the two contact devices, Dev01 and Dev02, and is nearly matched by Dev03, a stationary contactless device, as seen in Figure 20. Remarkably, smartphone application, Dev08, improves considerably over its single finger performance, notably with the LE matcher. Figure 11 and Figure 12 verify that single finger match scores tend to be among the highest for fingers 02, 03 and 07, 08 for all devices.



Figure 20 - Two + Two finger combination appears to produce the best performance for several devices, though with some disparity between the two matchers.

4.2.10 Six/Eight Fingers

Using all eight fingers (or six with the COM matcher, as explained in section 4.2.7) provides advantage for some DUTs, although the two + two combination is better for others, such as Dev08. As observed in [1], four-finger captures with contactless devices are challenged by geometric misalignment, lighting, and limited depth of field of the devices, especially smartphones.



Figure 21 – Eight finger combination generates lower error for most devices using the LE matcher, shown beside the six-finger combination, which generates lower error for most devices using COM matcher

4.2.11 Eight Fingers with Dev 03 As Exemplar (LE Matcher Only)

Dev03 emerged as the best performing of the contactless devices, so match error was examined with both contact and contactless probes matched to the Dev03 contactless fingerprints as exemplars using only the LE matcher for this experiment.



Figure 22 – FNIR with Dev03 prints as exemplars probed by prints from other devices, using LE matcher only.

Figure 22 shows some reduction in FNIR for Dev04 and Dev08. One possible cause for this may be that such devices perform better when matched against rolled exemplars. We have noted, while monitoring display of images during the registration process, that out-of-plane rotation of some fingers reduces the overlap such fingers might have with plain contact exemplars. Dev03 offers a somewhat larger sample of friction ridge that might increase match scores for these rotated contactless print samples. Another possibility is that contactless fingerprint impressions more closely resemble other contactless fingerprint impressions than they do exemplars from contact devices. The best test to determine the cause of the observed performance in Figure 22 would be to examine contactless print matching against rolled contact impressions – a notable goal for future data collection efforts.

4.3 Match Error Performance Summaries

The performance data presented in graphical form above is provided in tabular form in this section. Each table displays the error at one of two of the selected operating thresholds (LE matcher only) for selected finger combinations. First, False Negative Identification Rate (FNIR) is converted to True Positive Identification Rate (TPIR) in percent via the formula

$$TPIR(\%) = (1 - FNIR)*100$$
 (1)

Note that FPIR values are adjusted such that they are no lower than the upper limit of 95 % confidence interval for a binomial experiment consisting of *N* trials. By the Rule of Three, this interval is 0 to 3/N [6][7].

	Single Finger		Single Finger						Fing	ers	Fing	ers	Fing	ers		
					ngle Finger Fingers 02, 03		Fingers 02, 07		02,03,07,08		02,03,04,05		07,08,09,10		Eight Fingers	
	TPIR	FPIR ¹²	TPIR	FPIR ¹²	TPIR	FPIR ¹²	TPIR	\mathbf{FPIR}^{12}	TPIR	\mathbf{FPIR}^{12}	TPIR	FPIR ¹²	TPIR	\mathbf{FPIR}^{12}		
Dev01	91.94	0.19	97.94	1.55	98.97	1.55	99.48	1.55	98.45	1.55	98.97	1.55	99.48	1.55		
Dev02	82.60	0.19	96.41	1.54	98.46	1.54	99.49	1.53	98.47	1.53	97.96	1.53	99.49	1.53		
Dev03	71.11	0.19	92.46	1.51	91.46	1.51	95.98	1.51	93.47	1.51	94.47	1.51	97.49	1.51		
Dev04	52.16	0.19	83.76	1.52	74.75	1.52	82.83	1.52	87.31	1.52	80.61	1.53	88.89	1.52		
Dev05	26.52	0.19	59.30	1.51	53.00	1.50	68.00	1.50	68.84	1.51	62.31	1.51	80.00	1.50		
Dev06	24.76	0.19	63.92	1.55	69.59	1.55	70.62	1.55	67.01	1.55	67.01	1.55	76.29	1.55		
Dev07	7.98	0.19	23.86	1.52	33.33	1.52	30.30	1.52	34.01	1.52	36.87	1.52	50.00	1.52		
Dev08	35.23	0.19	76.77	1.52	81.31	1.52	83.84	1.52	80.30	1.52	75.25	1.52	83.33	1.52		

Table 2 – TPIR/FPIR at threshold 8 500 (%)

Table 3 – TPIR/FPIR at threshold 4 100 (%)

	Single Finger						Fing	ers	Fing	ers	Fing	ers		
			Single Finger		Ingle Finger Fingers 02, 03		Fingers 02, 07 02,03,07,08		07,08	02,03,	04,05	07,08,09,10		Eight Fingers
	TPIR	FPIR ¹²	TPIR	FPIR ¹²	TPIR	\mathbf{FPIR}^{12}	TPIR	\mathbf{FPIR}^{12}	TPIR	\mathbf{FPIR}^{12}	TPIR	FPIR ¹²	TPIR	\mathbf{FPIR}^{12}
Dev01	96.00	0.26	98.97	1.55	98.97	1.55	99.48	1.55	98.97	1.55	98.97	1.55	99.81	1.55
Dev02	91.49	0.19	97.95	1.54	98.97	1.54	99.49	1.53	98.47	1.53	98.98	1.53	99.49	1.53
Dev03	83.73	0.19	94.97	1.51	96.48	1.51	98.99	1.51	95.98	1.51	96.48	1.51	98.49	1.51
Dev04	68.83	0.25	86.80	1.52	84.34	1.52	89.39	1.52	92.89	1.52	88.78	1.53	91.92	1.52
Dev05	45.51	0.19	75.88	1.51	77.00	1.50	86.00	1.50	81.91	1.51	74.37	1.51	88.00	1.50
Dev06	45.84	0.19	76.80	1.55	77.84	1.55	78.87	1.55	76.80	1.55	78.35	1.55	78.35	1.55
Dev07	20.42	0.19	48.73	1.52	55.05	1.52	59.60	1.52	51.78	1.52	52.53	1.52	63.13	1.52
Dev08	53.28	0.19	86.87	1.52	90.91	1.52	92.42	1.52	87.88	1.52	85.35	1.52	89.39	1.52

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8307

¹² Note that grouping samples into finger combinations results in reducing the value of *N*, which increases the value of the *3/N* upper limit. As a result, this leads to an increase in the computed FPIR values shown in the Tables Table 2 and Table 3.

4.4 Matcher Tests: Area Under the ROC Curve

The current datasets are small in comparison to the hundreds of thousands of scores usually encountered in our biometrics testing [11]. As with data analyses in medical research often restricted to small samples, ROC analysis may be summarized effectively by the Area Under the (ROC) Curve (AUC) [12]-[16].

The AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [17]. If the trapezoidal rule is employed, the AUC is equivalent to the Mann-Whitney statistic formed by independent and identically distributed (i.i.d.) genuine scores and impostor scores [10]-[18].

Let $f_G(s)$ and $f_I(s)$ denote the continuous probability density functions (pdf) of the genuine scores and the impostor scores at a score $s \in \{s\}$, respectively. Then, the AUC can be expressed as

$$AUC = \int_{-\infty}^{+\infty} \left[\int_{s}^{+\infty} f_{G}(t) dt \right] \times f_{I}(s) ds .$$
(1)

Its derivation can be found in [18].

Let $P_i(s)$, where $s \in \{s\}$ and $i \in \{G, I\}$, denote the discrete empirical probabilities of genuine scores and impostor scores occurring at a score s, respectively. Then, by employing the trapezoidal rule, the AUC expressed in Eq. (1) can be estimated as follows:

$$\hat{A} = \sum_{s=s \max}^{s \min} \text{trapezoid (s)}$$

$$= \sum_{s=s \max}^{s \min} \text{triangle (s)} + \sum_{s=s \max}^{s \min} \text{rectangle (s)}$$

$$= \sum_{s=s \max}^{s \min} P_{I}(s) \times \left[\frac{1}{2} \times P_{G}(s) + \sum_{\tau=s+1}^{s \max} P_{G}(\tau)\right].$$
(2)

In the figures to follow, we display the AUC computed considering the genuine and impostor score distributions displayed as a point. The distributions are resampled with replacement to generate 2 000 replicate distributions and the computation repeated to form a distribution of AUC values from which we determine the upper and lower 95 % confidence limits on the AUC. The confidence interval thus determined by the bootstrap procedure is displayed around the initially computed value.

4.4.1 Single Finger

Examining Figure 23, we see that some devices exhibit slightly better performance using the COM matcher. As was pointed out previously the single finger scores include all fingers, including the little fingers.



Figure 23 – AUC for single finger from both matchers

4.4.2 Index Finger Only

Figure 24 and Figure 25 display AUC for single index fingers matched with LE and COM matchers, respectively. It is notable that with only a single index finger, the contact devices, Dev01 and Dev02, score very well. Dev08 also scores well at an AUC of 0.9 with only the single index finger.



Figure 24 – AUC for index finger from right hand using both matchers



Figure 25 – AUC for index finger from left hand using both matchers

4.4.3 Two Fingers

In Figure 26, with the addition of the middle finger to the index, the second stationary contactless device, Dev04, and smartphone, Dev08, are shown to approach or exceed the 90 % mark with Dev03 and the two contact devices.



Figure 26 – AUC for index and middle fingers from right hand using both matchers

The first two fingers of the left hand exhibit performance similar to that for the right hand except that Dev03 achieves a very high AUC with a tighter confidence interval with the COM matcher.



Figure 27 – AUC for index and middle fingers from left hand using both matchers

4.4.4 Three/Four Fingers

With additional sampled fingers of the right hand, further improvement can be seen with Dev04 breaking the 90 % mark with both matchers along with Dev08 for the LE matcher.



Figure 28 – AUC for four fingers from right hand using LE matcher shown beside AUC for first three fingers using COM matcher

A similar pattern is observed for the fingers of the left hand with near-perfect AUC scores for the contact devices and a score above 0.99 for the stationary contactless device, Dev03, with the COM matcher



Figure 29 – AUC for four fingers from left hand using LE matcher shown beside AUC for first three fingers from left hand using COM matcher

4.4.5 One + One Index Finger Combination

As shown in Figure 30, with two index fingers, the two contact devices display AUC values at or exceeding 98 %, Dev03 almost reaching 98 %, and one of the mobile contactless devices, Dev08, almost at 94 %.



Figure 30 – AUC for index fingers from both hands using both matchers

4.4.6 Two + Two Finger Combination

Shown in Figure 31, with the index and middle fingers of both hands, Dev03 reaches 99 % for both matchers. Dev08 is at 94 % with Dev04 and mobile device Dev05 breaking the 90 % mark.



Figure 31 – AUC for index and middle fingers from both hands using both matchers

4.4.7 Six/Eight Fingers

It can be observed in Figure 32, when using all available fingers, the contact devices, Dev01 and Dev02 are over 99 % with tight confidence intervals.



Figure 32 - AUC for all eight fingers from both hands using LE matcher shown beside AUC for the first three fingers from both hands using COM matcher

4.4.8 AUC Results by Device

In the following sections, we examine AUC for the various finger combinations for each device.

4.4.8.1 Dev01

The control contact device shows near perfect performance for all but the single finger (all finger) matching. It is worth noting that both exemplar and probe samples are acquired using the same Appendix F certified device for all Dev01 results. Figure 33 illustrates AUC values \geq 99 % for most finger combinations, though the confidence intervals are broader for the COM matcher, even as the AUC is virtually 100 % for all finger combinations.



4.4.8.2 Dev02

Here, probe samples acquired using an Appendix F device that uses a different imaging technology than that used to sample exemplar images are evaluated. As seen in Figure 34, performance remains high, but not quite reaching the level observed when both probe and exemplar are acquired using the same device (i.e. Dev01). However, in evaluating contactless fingerprint devices, it is important to acknowledge that some variation in performance is to be expected even among Appendix F certified devices.



Figure 34 – AUC for all finger combinations with Dev02 using both matchers

4.4.8.3 Dev03

Figure 35 indicates that the Dev03 contactless device shows AUC at or near 99 % with some multiple finger combinations.



Figure 35 – AUC for all finger combinations with Dev03 using both matchers

4.4.8.4 Dev04

Figure 36 shows that Dev04 approaches or exceeds the 95 % mark with some multiple finger combinations.



Figure 36 – AUC for all finger combinations with Dev04 using both matchers

4.4.8.5 Dev05

Figure 37 shows a few finger combinations that reach the 90 % AUC with the LE matcher. Interesting, but unexplained, is the much lower performance observed with the COM matcher. We suspect that this particular matcher may rely heavily on detail of the core region more so than the LE matcher.



4.4.8.6 Dev06

Figure 38 illustrates the disparity in performance with the two different matchers for Dev06.



Figure 38 – AUC for all finger combinations with Dev06 using both matchers

4.4.8.7 Dev07

Dev07 exhibits anomalously poor performance across all finger combinations (see Figure 39) yet showing its best performance with the "two + two" finger combination found to perform well for all devices.



4.4.8.8 Dev08

Figure 40 illustrates that Dev08 surpasses the 90 % mark with several finger combinations, with highest AUC with the "two + two" combination.



Figure 40 – AUC for all finger combinations with Dev08 using both matchers

5 Discussion and Conclusions

5.1 Metrics

A few of the comparison metrics detailed in [1] were applied to fingerprint images output from the devices under test (DUTs) and corresponding exemplars captured with the control FTIR device. We see that control of scale or sample rate continues to present a challenge for mobile contactless capture, though it appears to be much improved for stationary contactless capture, even to the extent that one stationary device has its entire distribution of measurements within the PIV range of 500 ppi ± 10 ppi. Other contactless devices show a little over 50 % within the PIV range with medians at 500 ppi, however exhibiting larger dispersion. The matchers used in the current study expect images sampled at 500 ppi, hence the observed match performance correlates roughly with the order of the scale and several other metrics. Regarding variations in scale and minutiae displacement, it is worth noting that modern matchers demonstrate some, even if limited, invariance to rotation and robustness to limited scale variation or displacement of minutiae as observed with deformation of the skin during contact capture. The matchers used here demonstrate some of this robustness but do expect either explicit input of the sample rate of fingerprint images or that the input images are conformant to the default sample rate of 500 ppi.

5.2 Matching

One of the principal objectives of the current research effort is to verify that contactless matching is much improved when considering multiple fingers. This hypothesis is confirmed. Results show that all fingers are not created equal from the point of view of matching and that, while treating each finger of *N* subjects as independent might be attractive to increase sample size, the practice is not realistic from an operational point of view. Images captured from index and middle fingers of either hand appear to carry the most information to support matching for both contact and contactless devices. These two fingers of either hand, or preferably of both hands, show substantial improvement over most other combinations. For contactless devices, particularly smartphone applications, optimizing for focus and lighting of index and middle fingers should produce the best results. As we observed in [1] and confirmed via extensive interaction with the smartphone applications, maintaining focus over four fingers simultaneously is difficult given the small depth of field of smartphone cameras.

5.3 Other Conclusions and Observations

We have observed a trend across most metrics where the performance of the DUTs appear to fall into three tiers of performance. The first, highest performing tier unsurprisingly includes the two contact devices, followed by a second, slightly lower-performing tier consisting of the two stationary contactless devices, and finally a third tier consisting of the mobile smartphone contactless devices. This trend is not surprising in that it is logical to expect that measurements of interoperability will show lower performance when comparing samples collected from devices with increasing deviation from the capture methods and technology employed by the control device. This is borne out in the results seen in comparing the two contact devices, which generally show only small differences, as opposed to the stationary contactless devices which show greater difference followed by the mobile contactless devices, which show even greater difference.

6 References

- [1] Libert J, Grantham J, Bandini B, Wood S, Garris M, Ko K, Byers F, Watson C (2018) Guidance for Evaluating Contactless Fingerprint Acquisition Devices, NIST Special Publication 500–305, National Institute of Standards and Technology, Gaithersburg, MD.
- [2] FBI/CJIS, "Electronic Biometric Transmission Specification (EBTS) Technical and Operational Update (TOU) 10.0.9 Redlines", Federal Bureau of Information Criminal Justice Information Systems, Clarksburg, WV, May 22, 2018,

https://www.fbibiospecs.cjis.gov/Document/Get?fileName=EBTS%20TOU%2010_0_9_Final.pdf Retrieved 05/28/2019

- [3] Personal Identity Verification (PIV) Image Quality Specifications for Single Finger Capture Devices, FBI/CJIS Biometric Specifications, 10 July 2006, <u>https://www.fbibiospecs.cjis.gov/Document/Get?fileName=pivspec.pdf</u>. Retrieved 04/18/2019.
- [4] Special Publication (NIST SP) 500-290e3: "American National Standard for Information Systems Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information, ANSI/NIST-ITL 1-2011 NIST Special Publication 500-290 Edition 3. August 22, 2016. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-290e3.pdf
- [5] Minutiae Interoperability Exchange (MINEX), NIST, <u>https://www.nist.gov/programs-projects/minutiae-interoperability-exchange-minex-overview</u>.
- [6] Hanley, JA; Lippman-Hand A (1983). "If nothing goes wrong, is everything alright?". JAMA 249 (13): 1743-5. PMID 6827763
- [7] Jovanoviv, B. D., and P. S. Levy. A look at the Rule of Three. The American Statistician, Vol. 51, No. 2, May 1997, pp. 137-139.
- [8] Maltoni D, Maio D, Jain A K, Prabhakar S. Handbook of Fingerprint Recognition. Springer, 2009, pp. 17
- [9] J.C. Wu, A.F. Martin, and R.N. Kacker, "Monte Carlo studies of bootstrap variability in ROC analysis with data dependency," Communications in Statistics – Simulation and Computation, vol. 48, no. 2, pp. 317-333, 2019
- [10] J.C. Wu and C.L. Wilson, "Nonparametric analysis of fingerprint data on large data sets," Pattern Recognition, vol. 40, no. 9, pp. 2574-2584, 2007.
- [11] J.C. Wu and C.L. Wilson, "An empirical study of sample size in ROC-curve analysis of fingerprint data," in Biometric Technology for Human Identification III, Proceedings of SPIE Vol. 6202, 620207, 2006
- [12] J.A. Hanley and B.J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," Radiology, vol. 143, pp. 29–36, 1982.
- [13] J. A. Hanley and B.J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," Radiology, vol. 148, pp. 839-843, 1983.
- [14] D. Bamber, "The area above the ordinal dominance graph and the area below the receiver operating characteristic graph," J. Math Psych, vol. 12, pp. 387-415, 1975.
- [15] G.E. Noether, Elements of nonparametric statistics. pp. 31-32. John Wiley and Sons, New York, 1967.
- [16] E.R. DeLong, D.M., DeLong, and D.L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," Biometrics, vol. 44, pp. 837-845, 1988.
- [17] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, pp. 861–874, 2006.
- [18] J.C. Wu, A.F. Martin, and R.N. Kacker, "Validation of nonparametric two-sample bootstrap in ROC analysis on large datasets," Communications in Statistics – Simulation and Computation, vol. 45, no. 5, pp. 1689-1703, 2016.