Minutiae Interoperability

Elham Tabassi, Patrick Grother, Wayne Salamon, and Craig Watson

{tabassi,pgrother,wsalamon,cwatson}@nist.gov

Abstract: Many large scale identity management applications require storage and exchange of standardized minutiae templates. Minutia templates offer a more space-efficient, less resource intensive, and more cost effective alternative to raw images. Recent minutiae interoperability tests (ILO, MTIT, MINEX) all reported variation in minutia selection and placement as the major factor affecting interoperability. This paper quantifies their effects and investigates how variation in selection and placement of minutia from different suppliers relates to loss of performance compared with proprietary templates. We concur with MTIT findings that conformance testing methodologies for evaluating the semantic content of minutia templates is essential and interoperability can be improved by closer adherence to the minutia placement requirement defined in a standard.

1 Introduction

Use of fingerprint templates is increasingly favored over the use of conventional fingerprint images mostly due to its compact representation, and also for privacy concerns. A fingerprint image requires a considerable amount of memory for storage (about 200 Kbytes uncompressed and 15 Kbytes compressed), as opposed to fingerprint templates that are only a fraction of that size (about 300 bytes). Also, the use of fingerprint templates are believed to be more secure allowing privacy sensitive solutions. Addressing size and privacy concerns, a more compact representation of fingerprint images, or templates, has gain acceptance as an alternative to the use and exchange of images for fingerprint matching in dissimilar applications.

A template is a list of specific friction ridge characteristics from a fingerprint image. Minutiae points are local ridge characteristics where a friction skin ridge begins, terminates, or splits into two or more ridges. A minutia point is generally described by its position and orientation in a fingerprint. For many applications, minutiae templates offer a more space-efficient, less resource intensive, and more cost effective alternative to raw images.

For open systems use of minutiae templates as the medium for fingerprint interchange may adversely affect the interoperability and hence performance. Different vendors use different coordinate systems, location and angle definitions to describe the same minutia. These differences could result in lower accuracy of fingerprint matching systems that exchange minutiae extracted using different methods rather than exchange of finger images. Consequently, to improve interoperability, standards have been developed to specify the location and formatting of minutiae data, (i.e. minutiae template), for matching purposes

[JTC08, Ame08, MN07]. These standards create the possibility of a fully interoperable multivendor marketplace for applications involving fast, economic, and accurate interchange of compact biometric templates. To assess the sufficiency and performance of these standards, several evaluations [$G^{+}06$, Int05, UK 06] have been organized to quantify interoperability and performance degradation of fingerprint matching systems using standard templates compared with proprietary templates.

This paper reviews the problem of interoperability identified in the recent tests and focuses on the factors associated with degraded interoperability when minutiae templates are exchanged. Section 2 gives an overview of the existing minutiae standards. Section 3 reviews federated applications that require interoperable subsystems. The objective of interoperability tests is listed in Section 4 which is followed by overview of NIST Minutia Exchange Interoperability Test and its findings in Section 5. That gives context for our examining of causes of loss in performance when using standard minutia templates vs. proprietary image-based templates in Section 6 which is the main focus of this paper, followed by conclusions and way forward in Section 7.

2 Minutiae standard templates

The first minutiae standard was established in 1986 when the Federal Bureau of Investigation and National Institute of Standards and Technology (formerly the National Bureau of Standards) developed the minutiae-based ANSI/NBS-ICST 1-1986 Data format for fingerprint information interchange standard [McC04]. The standard has been revised three times since, but its latest version; ANSI/NIST-ITL 1-2007 Type-9 Record [MN07]; includes many of the requirements from its original standard. ANSI/NIST Type-9 minutiae information may be extracted and encoded in any of several different manners depending on the system that is used to scan an image, extract minutiae, and encode the minutiae template. The "standard format" defines a common block of tagged fields including mandatory minutia location, angle, type (ridge ending, bifurcation, compound, and undetermined), quality, finger position, finger pattern classification that produced minutia information, and optional data such as ridge count data and core or delta information. Additional reserved blocks are registered and allocated for use by specific vendors allowing them to encode minutiae data and any additional required characteristic or feature data in accordance with their own systems specific hardware and software configuration.

Developed in 2004 and currently under revision, the INCITS 378-2004 Fingerprint minutia format for data interchange [Ame08] is driven by commercial verification rather than law-enforcement identication needs. This standard was based on the ANSI/NIST-ITL1-2000 standard and the FBIs electronic fingerprint transmission specication (EFTS 7.0). The standard specifies how to compute minutia location and angle. Minutia type and quality are also recorded. Unlike ANSI/NIST-ITL1-2007 that uses lower left of an image as the origin, this format uses the upper left corner of the image. A minutia's angle is stated in increments of two degrees. The standard also has provision for an open format defined for the optional inclusion of common extended data fields. These include core and delta information, ridge count information for either four-neighbor quadrants or eight-neighbor octants, and vendor-defined information. The INCITS 378-2004 also contains provision for formatting data from several presentations or views of the same finger thus accommodating systems that rely on several readings of the same finger to construct a good average template.

International standard ISO/IEC 19794-2 Information technology-Biometric data interchange formats-Part 2: Finger minutia data [JTC08] was developed in 2005 and is currently under revision. Its structure is quite similar to the INCITS 378-2004 standard. The most significant difference between the ISO standard and the INCITS 378-2004 is the representation of minutiae angle which is 2 degrees increments in INCITS 378-2004 as opposed to 1.40625 degrees in the ISO version. As different vendors quantize to different values before mapping to 2 degree increments, this change in representation may not be signicant. ISO/IEC 19794-2 also defines compact representation of minutiae data for storage on smart cards.

3 Interoperable federated applications

Interoperability is not always a requirement for biometric systems, but only when the sources of its different subsystems are different suppliers. Generally speaking, a biometric system is a combination of several subsystems: data acquisition subsystem; transmission and data storage subsystem; template generation (or feature extraction) subsystem for subsequent comparison against stored templates; and finally decision making (or matching) subsystem based on comparison scores, thresholds and possibly other information like biographical or fusion information.

For closed systems, when the supplier of the different subsystems is the same, there is no interoperability issue. Otherwise, high performance would be achieved only if the various subsystems could successfully interoperate. Large-scale identity management applications such as personal identity verification (PIV)[Com], transportation security agency transportation worker identification credential (TWIC) program, and registered travelers (RT)[tsa08] in the U.S. as well as European citizen card are example of large-scale biometric systems that interoperability of its subsystems is essential. In the context of minutia interoperability, that means that minutia extractor algorithm and minutia comparison algorithm of a biometric system should be interoperable. Figure 1 shows the most general scenario for minutia interoperability. Fingerprint images are acquired using capture device A at the enrollment where enrolled templates are generated using algorithm X. Capture device B and template generator Y are used for authentication. Finally minutia comparison algorithm Z, compares minutia templates generated by algorithms X and Y. This is three-way interoperability because algorithms X, Y, and Z need to interoperate. However, often the minutia extractor and matcher of authentication phase (i.e. algorithms Y and Z of Figure 1) are from the same supplier, which makes it a two-way interoperability problem instead of three-way.

If template data rather than fingerprint image could be used with sufficient accuracy in a multi-vendor system, then bandwidth, storage space, and number of template extractions would all be substantially reduced.

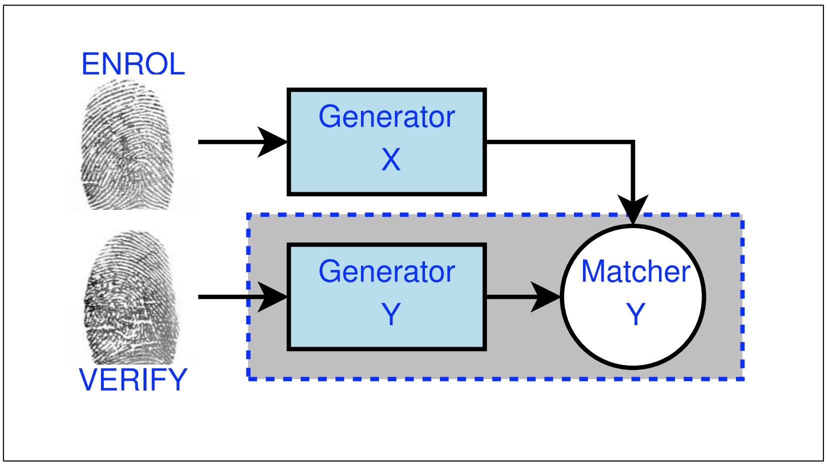


Figure 1: Three-way interoperability: Fingerprint images are acquired using capture device A at the enrollment where enrolled templates are generated using algorithm X. Capture device B and template generator Y are used for authentication. Finally minutia comparison algorithm Z, compares minutia templates generated by algorithms X and Y.

4 Interoperability tests

With increasing number of applications built on standardized templates questions arise regarding interoperability and sufficiency (performance) of the data interchange standard. A data interchnage format is sufficient when information coded in a standrad template is sufficient to enable successful recognition. In other words, error rates when comparing standardized templates are comparable with that of image-based proprietry templates of a leading minutiae extractor algorithm. This is distinct from the issue of interoperability, which mainly considers whether the comparison subsystem is able to process templates generated by different minutia extraction algorithms. Therefore, there is two layers to any interoperability test; interoprability and sufficiency. Minutiae template interoperability testing (MTIT) [UK 06] and the international standard on biometric performance testing and reporting (ISO/IEC 19795-1) [JTC05] refers to these as basic-interoperability and perfomance-based interoperability.

There have been several fingerprint recognition evalutions; fingerprint vendor technology evaluation (FpVTE)[WH$^+$04], fingerprint verification competition (FVC)[oBUUdM06], ongoing NIST proprietary fingerprint templates evaluations (PFT)[W$^+$08] to name a few, but there have been very few interoperability evaluations. The first interoperability test was performed in 2003 by the international labour organization (ILO)[Int05]. NIST initiated minutia exchange interoperability test (MINEX)[G$^+$06] in 2004 which currently is an on-going evaluation. Minutia template interoperability testing (MTIT) was performed by UK national physics laboratory in 2005. These tests were designed to determine and improve the feasibility of using standard minutiae templates as the interchange medium for fingerprint information between dissimilar fingerprint matching systems. These tests unanimously reported minutia selection and placement as main factors affecting interoperability, without quantifying their effect. This paper aims to investigate how variation in selection and placement of minutia from different suppliers relates to loss of performance compared with proprietary templates using MINEX data. An overview of MINEX follows.

5 The minutia exchange interoperability (MINEX) test

In 2004, national institute of standard and technology performed a large scale minutiae interoperability test to evaluate a) interoperability of the two minutiae extraction subsystems that generate standardized INCITS 378-2004 templates with respect to a comparison subsystem; and b) whether use of standardized minutiae templates instead of image data would result in successful match, i.e. if use of INCITS 378-2004 minutae template as opposed to image results in comparable error rates. The former evaluates feasibility of INCITS 378-2004 minutia templates and the latter its sufficiency. [JTC05] regards these two as basic interoperability and performance-based interoperability.

MINEX is by some measures the largest biometric test ever conducted. It involved testing the core template handling competency of fourteen fingerprint vendors using fingerprint images from a quarter of a million people, and executing in excess of 4.4 billion comparisons, in the production of more than 23,000 detection error tradeoff (DET) characteristics from 493418 mate (same-person) and 975890 non-mate (different person) comparisons.

5.1 Test design

MINEX test design is explained in [G^{+}04]. Each vendor participant provided NIST with their SDK that contained binary C libraries to:

1. create an INCITS 378-2004 MIN:A template from an image, coding minutia location (x, y), angle (θ), and type,
2. create an INCITS 378-2004 MIN:B template from an image, MIN:B is MIN:A with additional ridge count, core and delta information (this was optional),
3. create a proprietary template from an image,
4. produce a comparison score from two MIN:A templates,
5. produce a comparison score from two MIN:B templates (optional), and
6. produce a comparison score from two proprietary templates.

The minutiae quality field required by INCITS 378-2004 was set to zero in all cases, as no universally accepted denition for it exists. Creation of MIN:B templates were optional and only six out of fourteen vendors supplied MIN:B .

To establish a baseline set of performance statistics, MINEX participants were required to generate and compare the proprietary minutiae templates using their propriery minutia extraction and comparison algorithms.

In addition to the proprietary template generation and comparison functions, each MINEX vendors SDK was required to encode and compare MIN:A templates. In these "native" comparions, minutae template representation is constrained by the INCITS 378-2004 specifications while there is no constrain on "proprietary" comparisons. Therefore, the "properiety" comparisons are expected to give better performance than using MIN:A or MIN:B templates. Sufficiency of INCITS 378-2004 was quantified by the performance loss of properietry vs native comparions.

MINEX considered the two-way and three-way interoperability scenraios. Specifically four scenaroios were examined:

1. Enrollment template is generated with supplier X and compared with a template generated with supplier Y in verification transation using comparison algorithm of supplier Y. This is a two-way interoperability and reflects the typical access control situation in which supplier Y's generator and comparison algorithm are bundled together.

2. Compariosn algorithm Z compares templates generated by supplier X and supplier Y. This three-way interoperability scenario (as shown in Figure 1) is the most general case.

3. Comparison algorithm Z compares templates generated by the same supplier (X). This is commercially atypical but was included to examine whether comparison algorithm's dealing with the same-kind templates could result in any performance gain.

4. Comparison algorithm Z compares templates generated by supplier X and supplier Y from the same image. This examines the core of interoperability failure when effect of any difference in image due to re-capture are isolated.

MINEX used the false-non-match-rate (FNMR) at a fixed false-match-rate (FMR) as the figure of merit. The FNMR is the fraction of same-person comparisons that result in a comparison score less than or equal to the operating threshold of the comparison subsystem. FNMR is a measure of inconvenience i.e. the fraction of genuine transactions that result in failure. Likewise the FMR is the fraction of non-mate comparisons that result in a comparison score greater than the operating threshold. FMR is regarded as a measure of security, i.e. the fraction of illegitmate matching attempts that result in success. As is typical in offline testing [JTC05], MINEX did not fix an operating threshold but instead uses all the scores from a comparison algorithm as thresholds that could be used in actual operation. This contrasts with scenario testing which often uses a device configured with one fixed operating threshold. The output is then a decision and not a score, and this precludes investigation of performance at other thresholds.

5.2 Goals

MINEX objectives, as stated in [G+06], were to assess the viability of the INCITS 378-2004 [Ame08] templates as the interchange medium for fingerprint data. Three specific objectives were

1. To determine whether standardized minutiae enrollment templates can be subsequently matched against an authentication template from another vendor,
2. To estimate the verication accuracy when INCITS 378-2004 templates are compared relative to existing proprietary formats, and
3. To compare the INCITS 378-2004 template enhanced with ridge count “extended“ data (MIN:B) with the standards base template (MIN:A).

The first item is the interoperability test and measures core capability of comparison algorithms to process INCITS 378-2004 tempaltes generated by different minutia extraction algorithms. The second item is the sufficiency test and measures performance loss of using INCITS 378-2004 templates instead of image-based proprietry templates. The last item examines the utility of additional ridge count, core and delta information in the extended data fields of INCITS 378-2004 and if it could improve performance.

5.3 Datasets

Four datasets were used in MINEX testing that represented a range of operational image qualities. All of these are operational data sets gathered in on-going US Government operations, and have been sequestered at NIST for testing. MINEX uses randomly selected extracts of those databases. The integrity of the ground truth of the datasets was assured by human inspection. The quality composition of the datasets is tabulated using the NIST fingerprint image quality (NFIQ [TWW04, TW05]) method in Table 1. NFIQ summarization is performed according to recommendations in [TG04].

All datasets used were left and right index fingers only using live-scan plain impressions. The original images were given to NIST already WSQ compressed at approximately 15:1. The images were given to the template extraction algorithms as decompressed (using NISTs WSQ decoder) “raw“ pixel data. The original target sample sizes were 62,000 mates and 122,000 non-mates. These totals were reduced after consolidations and a few WSQ decompression failures were taken into account. The testing was performed by using the second instance of the mates as the enrollment image and the first instance as the authentication image. So for each dataset there were a little under 62,000 mate scores. The

Table 1: Summary of NIST fingerprint Image Quality values for the four MINEX data sets

Dataset	finger	1 Best	2	3	4	5 Worst	Summary
1	R	0.424	0.314	0.206	0.026	0.031	92.652
	L	0.442	0.268	0.212	0.034	0.044	91.089
2	R	0.437	0.338	0.157	0.007	0.061	90.912
	L	0.467	0.316	0.157	0.006	0.053	91.812
3	R	0.315	0.374	0.255	0.025	0.031	91.967
	L	0.348	0.3267	0.253	0.028	0.045	90.594
4	R	0.459	0.404	0.105	0.016	0.017	95.386
	L	0.432	0.375	0.143	0.021	0.029	95.386

non-mate scores were generated by comparing the non-mate authentication samples to the same enrollment images used with the mates, so for non-mate scores most enrollment images were used twice. This generated a little under 122,000 non-mate scores for a total of just under 184,000 scores per finger/dataset.

5.4 MINEX **findings**

As mentioned earlier, MINEX measured fingerprint matching error rates when multiple vendors generate and verify the interoperable templates standardized in INCITS 378-2004 . Specifically, MINEX evaluated a tripartite application paradigm in which the enrollment template, the verification template and the comparison algorithm could potentially be provided by different vendors. The study also compared performance available from standard templates with proprietary templates on the same datasets. Two- and three-way interoperability tests result in interoperability matrices of Table 2. The proprietary column shows performance figures (single finger false non-match rate at false match rate of 0.01) when both enrollment and verification templates are generated and compared using proprietary algorithms of a supplier. The native column shows performance numbers when INCITS 378-2004 templates are generated and compared with algorithms from the same supplier. Columns 4, 5, and 6 of Table 2 show, respectively, performance numbers when verification template generator and comparison algorithm are from the same supplier and different from the enrollment template generator, template generator and comparison algorithm are from different suppliers, and when template generator are from the same suppliers but different from the comparison algorithm supplier. Detail interoperability matrices are provided in [G$^+$06]. Qualitatively, the headline findings are that error rates

- are lowest when proprietary templates are used,
- increase when both templates and the matcher are from the supplier (native comparisons),

- increase further when both templates are generated by the same supplier but different from the comparison algorithm supplier, and
- are highest when template and matchers all come from different suppliers.

The loss in performance of proprietary systems compared with native comparisons is somehow expected since standard templates almost always contain less information than proprietary templates. The cost to achieve interoperability is that standard templates do not encode sufficient information needed to achieve performance level comparable with proprietary templates. The fact that using minutia template generators and comparison algorithms from different suppliers result in further performance loss points out the variations in selection and placement of minutiae by different extraction algorithms; either some minutiae are found by one algorithm and missed by the other one, or their encoding makes them look mismatch by the comparison algorithm. That suggests minutia extraction algorithms may systematically interpret a common input differently. The respective algorithmic difficulties are as follows:

- *Selection* - Different implementations will embed different approaches to detection of true minutiae and rejection of false minutiae. This may include regional biases such as ignoring minutiae in the periphery.
- *Placement* - Different implementations will generally report different values for (x, y, θ) despite the qualitative requirements on placement given in INCITS 378-2004 , clause 5.

Examples of these are depicted in Figure 2. Only the two out of six shown minutia are detected and placed similarly by the two minutia extraction algorithms. Note that different colors denote different minutia type. The overall (negative) effect on error rates is shown in Figure 3. Performance of native comparisons (i.e. the same supplier generated and compared standard templates) is always superior to the interoperable comparisons (i.e. comparison of standard templates generated by different suppliers), and in most cases rather significantly.

Further examination of why and how these factors affect interoperability is discussed in the following section.

6 Causes of interoperability degradation

The two major recent interoperability tests, MINEX [G$^+$06] and MTIT [UK 06], identified detection of false minutia and inconsistency in placement of true minutia as two major issues impacting interoperability. This section aims to quantify the effect of selection and placement of minutiae on performance, which is the main contribution of this paper. Detailed discussion follow.

Table 2: Interoperability Matrix: False non-match rate at false match rate of 0.01 for single finger verification. The two-way interoperability values are average over 8 minutia extractor algorithms. The three-way interoperability values are averages over 64 minutia extractor algorithms.

	Proprietary	Native	2 way-interoperability (mean FNMR)	3 way-Interoperablility (mean FNMR)	2 way-Interoperablility (mean FNMR)
Matcher		Enrollment = X Verification=X Matcher = X	Enrollment = X Verification=Y Matcher = Y	Enrollment = X Verification=Y Matcher = Z	Enrollment = X Verification=X Matcher = Z
Vendor 1	0.0089	0.0136	0.0273	0.0268	0.018
Vendor 2	0.0189	0.0251	0.0388	0.0413	0.0260
Vendor 3	0.0225	0.0225	0.0351	0.0373	0.0247
Vendor 4	0.089	0.0140	0.0209	0.0315	0.0225
Vendor 5	0.0047	0.0129	0.0303	0.0283	0.0191

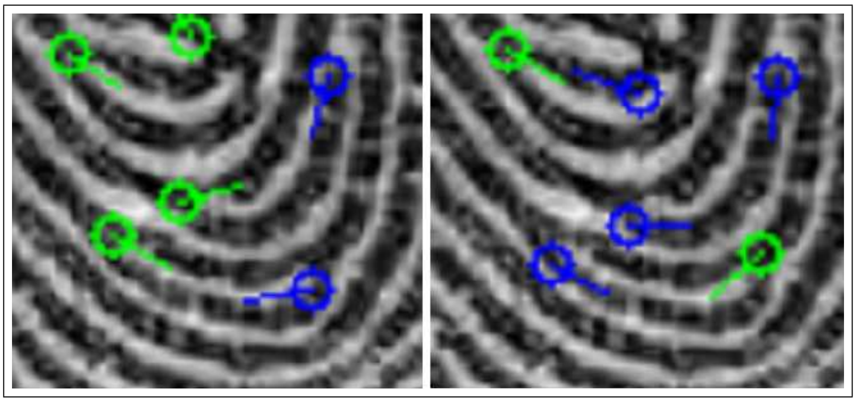


Figure 2: The results of alternative minutiae selection and placement algorithms: Note the angle differnce at top right, and the type, angle, and location differences at bottom right.

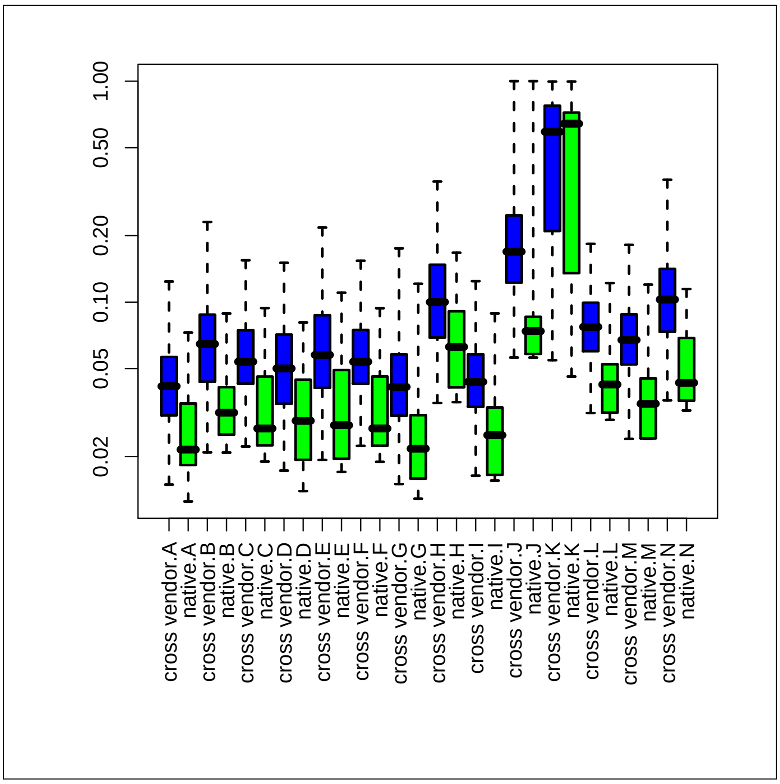


Figure 3: False non-match rate of native (standard minutia templates generated and compared with algorithms from the same supplier) and interoperable (standard minutia templates generation and comparison are performed using algorithms from different suppliers) at a false match rate of 0.01.

6.1 Effect of minutia placement on performance

The objective is to quantify the effect of variation in (x, y, θ) encoding by different minutia extractor algorithms on performance. To do so, we first calculated the number of minutia that are found to be the same within two finger minutia templates of the *same* image created by two different extractors. The criteria for overlapping minutiae is that the (x, y) coordinate of one minutia falls within radius R of an imaginary circle drawn about the second minutia's coordinate (We used $R = 5$ pixels). The fraction of overlapped minutiae is the size of the intersection set divided by the size of the smaller of the number of minutiae in the two input templates. Further, we calculated the mean displacement of those minutiae that are found to be paired as well as the difference in their angle.

Suppose M_i^n is the set of minutiae templates created by algorithm i from the n^{th} image ($n = 1 \dots N$). That is,

$$M_i^n = \{(x_k, y_k, \theta_k) \,|k = 1 \dots K_i\} \tag{1}$$

For each image $n = 1 \dots N$, the set of minutiae in common between extractors i and j is given by

$$R_{ij}^n = \{(k, l) \mid d_{ij}^n(k, l) \leq 5 \; and \; i \neq j\} \tag{2}$$

where

$$d_{ij}^n(k, l) = \sqrt{(x_k^n - x_l^n)^2 + (y_k^n - y_l^n)^2} \tag{3}$$

is the distance between the k^{th} minutia of extractor i and the l^{th} minutia of extractor j from the n^{th} image.

We picked the seven better performers of MINEX participants which result in $C(7, 3) = 210$ different combinations of minutia generators and matcher. To estimate variation in minutiae placement by different suppliers, we selected a random subset of $N = 20,000$ right index images of MINEX Dataset 1 (see Table 1). For each $n = 1 \dots N$ image, we computed the fraction of overlapped minutiae for algorithms i and j

$$m_{ij}^n = \frac{|R_{ij}^n|}{min(|R_i^n|,\; |R_j^n|)} \tag{4}$$

mean misplacement over all of overlapped minutia

$$d_{ij}^n = \frac{\sum_{(k,l)} d_{ij}^n(k, l)}{|R_{ij}^n|} \tag{5}$$

and mean angular difference of overlapped minutiae

$$A_{ij}^n = \frac{\sum_{(k,l)} |(\theta_k^n - \theta_l^n)|}{|R_{ij}^n|} \tag{6}$$

To get a summary statistic of the above three quantities for each (enrollment, verification) minutiae generator pair, we computed their 1-percentile value over all $20,000$ images.

$$\begin{aligned}
\text{angleDifference} &= \{\text{CDF}^{-1}_{A_{ij}}(0.01) \mid \forall i \neq j \text{ extractors}\} \\
\text{fractionInCommon} &= \{\text{CDF}^{-1}_{R_{ij}}(0.01) \mid \forall i \neq j \text{ extractors}\} \\
\text{misplacement} &= \{\text{CDF}^{-1}_{d_{ij}}(0.01) \mid \forall i \neq j \text{ extractors}\}
\end{aligned} \tag{7}$$

We used a model that is additive in fraction of overlapped minutia, misplacement of overlapped minutiae, and difference in angle of overlapped minutia (eq. 7) to describe the performance loss of native comparisons (template generators and comparison algorithm from same supplier) compared with interoperable comparisons (template generators and comparison algorithm from different supplier). Performance loss is expressed as the delta between false nonmatch rate of native comparisons (F_{kkk}) and false non-match rate of interoperable comparisons (F_{ijk}) when threshold was set at native comparisons' false match rate of 0.01.

$$\begin{aligned}
F^t_{ijk} - F^t_{kkk} &= \alpha + \beta_1 \text{ratio} + \beta_2 \text{misplacement} \\
&\quad + \beta_3 \text{angleDifference} + \epsilon
\end{aligned} \tag{8}$$

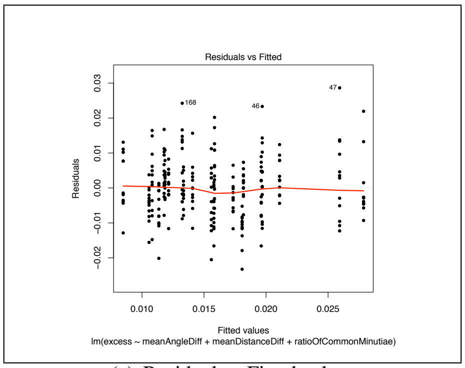
The result are shown in Table 3 and Figure 4. The residual error have an almost normal distribution which along with very small p-values suggest that all three factors are quite significant.

Table 3: Linear fit parameter of equation 8

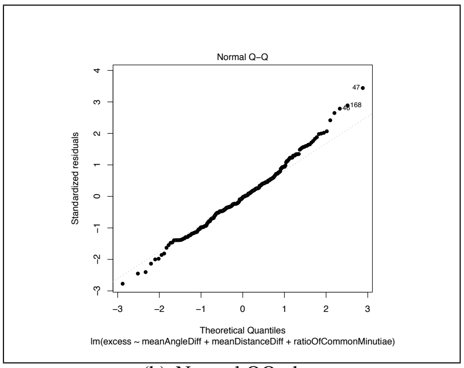
Coefficients	Estimate	Std. Error	t value	$Pr(>\|t\|)$
Intercept α	5.481968	0.600207	9.133	$< 2e-16$
mean angular difference β_3	-5.441200	0.597892	-9.101	$< 2e-16$
mean misplacement β_2	-0.013793	0.002778	-4.966	$1.28e-06$
fraction of overlapped minutia β_1	-0.033737	0.006147	-5.488	$1.00e-07$

6.2 Effect of minutia selection (or detection strategy) occurrence densities

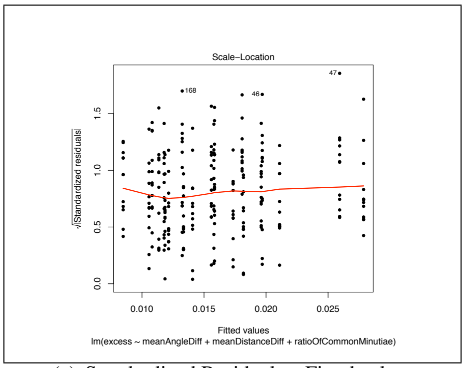
Consider a corpus of N single finger images collected in an operational scenario in which the right index finger of N subjects is stored as a greylevel raster of a fixed size. Suppose further that we apply a minutiae detection algorithm to each of those images and save the result as an INCITS 378 minutiae record. We then compute a two-dimensional histogram



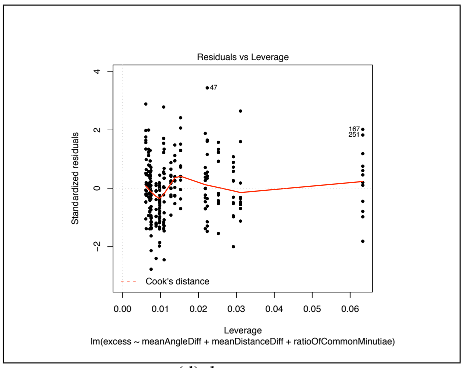
(a) Residual vs Fitted value



(b) Normal QQ plot

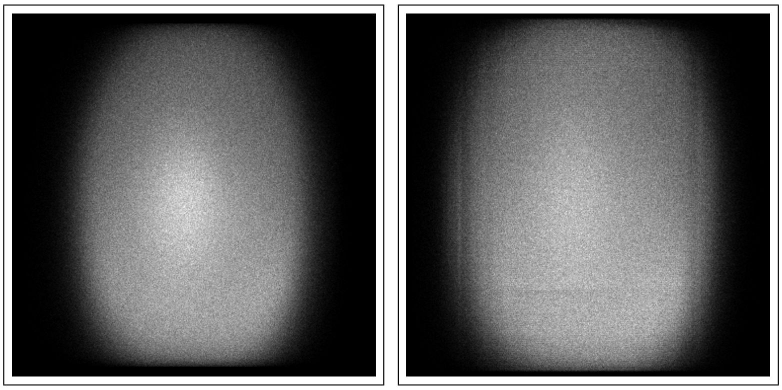


(c) Standardized Residual vs Fitted value



(d) leverage

Figure 4: Diagnosis plots of the linear fit of eq. 8. The almost normal distribution of residual error indicates that the loss of performance could mostly be explained by the three factors: fraction of overlapped minutiae, mean minutiae misplacement and mean angular difference of overlapped minutiae.



(a) Narrow (b) Wide

Figure 5: Examples of two-dimensional Minutiae Placement Density Functions.

of (x, y) location of minutia

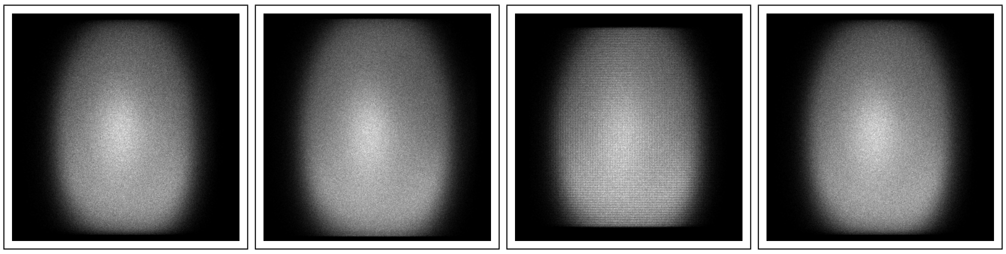
$$P(x, y) = \sum_{i=1}^{N} \sum_{k=1}^{K_i} \delta(x - x_{ik}, y - y_{ik}) \tag{9}$$

where the function $\delta_{ik}(x, y) = 1$ if and only if the k-th minutia of the i-th template is placed at position (x, y), and is zero otherwise.

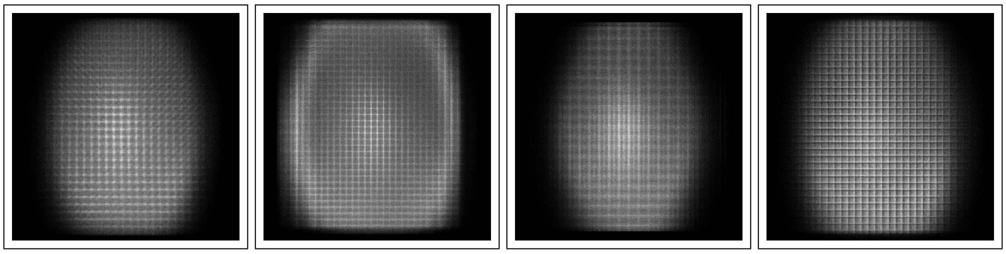
We computed such functions for the minutia extraction algorithms submitted to the MINEX evaluation, using a database of 368 by 368 optically acquired right index finger images from N = 183525 subject. We observed a median of $K = 38$ minutiae per record. Note that the input templates were not transformed in any way (by registration of the core, for example). The image population was not sampled by any factor (such as sex, quality, or image class), and the angle and type information for each minutiae was ignored.

The intent of this analysis was to discern whether some implementations exhibit regional preferences in how they find minutiae, for example whether some template generators center-weight while others weight in the periphery. Indeed such diversity is apparent in the images of Figure 5 which are min-max linearly scaled versions of the $P(x, y)$ estimates (eq. 9). Note that the template generator of Fig. 5(a) finds minutiae in a narrower region than that in Fig. 5(b). One notable caveat here is that these are systemic aggregated results and that for any given image (including perhaps the difficult-to-match ones) the behaviour of the two minutiae extractors may be very similar.

A second, unexpected, finding of this computation is that almost half of the MINEX minutiae extractors exhibit a periodic structure in their respective $P(x, y)$. Some of these are shown in the second row of Figure 6. Why this occurs is not known, except perhaps to the algorithm developers. One possible explanation would be the periodicity arises as an artifact of a Fourier transform operation applied to tiles across the image. Such an



(a) Smooth PDF



(b) Periodic structure in PDF

Figure 6: Examples of 2D Minutiae Placement Density Functions. In the top row the figures show more or less expected form; In the second row the minutiae densities exhibit a variety of periodic patterns. The order of appearance here does not correspond to the alphabetic ordering used in the MINEX report.

approach would not necessarily have to be incorrectly implemented because the periodicity would be a natural, and unintended, consequence of the operation.

The presence of fine grained periodic structure in a two-dimensional histogram such as Figure 6 is inconsistent with the expected smoothly varying form. Indeed it indicates that minutiae are not being placed in the natural locations. Obviously, this minutia misplacement adversly affects interoperability because higher performing template generators of the MINEX test do not exhibit such pattern. Furthermore, minutia extractor algorithms that exhibit such behavior are *probably* not conformant to the INCITS 378-2004 standard. We say probably and not definitely because in principle, a generator could chose to either preferentially include minutiae that actually fall on the grid, or preferentially exclude those that don't. In both cases, the result would be as shown in Figure 6. If such a strategy was implemented correctly then a conclusion of non-conformance with INCITS 378-2004 is incorrect. Note, however, there seems to be no good reason to implement such schemes as this would naturally reduce the number of minutiae.

7 Conclusion

The variation in minutiae selection and placement by different vendors is correlated with degraded interoperability. The 2D minutiae occurrence density functions (Figure 6) suggest that INCITS 378-2004 minutiae template generation is idiosyncratic. Specifically the periodic grid patterns indicate that the encoder is tending to quantize minutiae location, and so departing from the requirements of the standard. The exact behavior is different for each encoder, and is absent completely in some others. Such interoperability issues could be resolved by semantic conformance test which tests how faithful reported minutiae are to the underlying ground truth. Future studies include applying other methods such as generalized linear mixed model to quantify the effect of minutia placement and selection as well as advancing semantic testing methodologies.

References

[Ame08] American National Standard. ANSI-INCITS 378 Fingerprint Minutia Format for Data Interchange, August 2008.

[Com] Computer Security Division - NIST. Federal Information Processing Standards Publication 201 Personal Identity Verification for Federal Employees and Contractors.

[G+04] Patrick Grother et al. Minutiae Interoperability Exchange Test 2004 (MINEX04) API Specification, December 2004.

[G+06] Patrick Grother et al. MINEX Performance and Interoperability of the INCITS 378 Fingerprint Template, NISTIR 7296, April 2006.

[Int05] Geneva International Labour Organization. Seafarers' Identity Documents Biometric Testing Campaign Report, 2005.

[JTC05] JTC1 SC37 Working Group 5. ISO/IEC 19795 Biometric Performance Testing and Reporting: Part 1 - Principles and Framework, August 2005.

[JTC08] JTC1 SC37 Working Group 3. International Standard ISO/IEC 19794-2 Information Technology - Biometric Data Interchange Formats - Part 2: Finger minutia data, August 2008.

[McC04] R. Michael McCabe. *Automatic Fingerprint Recognition Systems*. Springer, New York, 2004.

[MN07] R Michael McCabe and Elaine M Newton. Data format for the interchange of fingerprint, facial and other biometric information - part 1, April 2007.

[oBUUdM06] Biometric System Laboratory (University of Bologna), Pattern Recognition Image Processing Laboratory (Michigan State University), Biometric Test Center (San Jose State University), and Biometrics Research Lab ATVS (Universidad Autonoma de Madrid). Fingerprint Vendor Competition, March 2006.

[TG04] Elham Tabassi and Patrick Grother. Quality Summarization - Recommendations on Biometric Quality Summarization across the Application Domain NISTIR 7422, August 2004.

[tsa08] TSA Registered Traveler Security, Privacy and Compliance Standard for Sponsoring Entities and Service Providers, January 2008.

[TW05] Elham Tabassi and Charles Wilson. A Novel Approach to Fingerprint Image Quality. In *Proc. IEEE ICIP'05*, volume 2, pages 37–40, Genova, Italy, September 2005.

[TWW04] Elham Tabassi, Charlse L. Wilson, and Craig I. Watson. NIST Fingerprint Image Quality NISTIR 7151, August 2004.

[UK 06] UK National Physics Laboratory. Minutia Template Interoperability Testing, June 2006.

[W$^+$08] Craig I Watson et al. NIST Proprietary fingerprint template (PFT) Testing, May 2008.

[WH$^+$04] Charles L. Wilson, Austin Hicklin, et al. Fingerprint Vendor Technology Evaluation NISTIR 7123, June 2004.