Interoperable Performance

Patrick Grother[1]

National Institute of Standards and Technology, MD 20899, USA
pgrother@nist.gov

Synonyms

None.

Definition

Accuracy available from a biometric system that includes standardized components from several suppliers.

In applications where components conform to standardized interfaces and functional specifications, it is possible to replace one component with another from a different manufacturer. While conformance to the specifications is a necessary condition for interoperability, it is often not sufficient because the internal algorithmic action of the component is usually not regulated by the standard. Thus a biometric detection algorithm might underperform some others despite being conformant to the requirements. This article suggests that the appropriate means for quantifying biometric interoperability is to identify relevant performance metrics, to measure them, and to certify against them. For biometric sensors and also detection, segmentation, and matching algorithms, these metrics will be usually be failure to acquire and enroll rates, and Type I and II recognition error rates.

Introduction

Biometric recognition is explicitly a two-phase operation: In verification, a first-encounter enrollment sample is compared with a second-encounter verification sample. Similarly in identification, a new sample is searched against a set of prior enrollments. If the samples are not captured and processed using the same hardware and software, identically configured, the issue of whether the various components are interoperable arises. While interoperability is a desirable and necessary aspect of applications in which multiple vendors sell equipment for capture, processing, and matching it rests on the availability of well crafted standards, and specifically conformance of the various components to those standards. So sensors might have to conform to imaging specifications, their outputs to image exchange standards, and their transmission might require equipment implementing standardized interfaces. The hazard in biometric applications is that a weak specification, or lack of conformance to a specification, might undermine the accuracy of the whole recognition system.

Figure 1 depicts a general interoperable application. It shows N different biometric capture devices (BCDs) being used to acquire sample data that is then converted from its raw captured biometric data block (cBDB) format into a standardized biometric data block (sBDB) format for enrollment. This is done by any of I template generators. Later, these will be compared by any of K comparison subsystems against verification (or identification) records (sBDBs, in this case) produced from any of J generators processing the output of M BCDs.

This formalism is notional; it defines a five dimensional component space the last of which is the comparison engine whose outputs support measurements of accuracy. Thus any combination of five different products can be tested. An interoperable component is then one that can be used in combination with others. Note that this defines biometric interoperability differently than in some other domains where strict conformance guarantees performance. For example, while a non-conformant

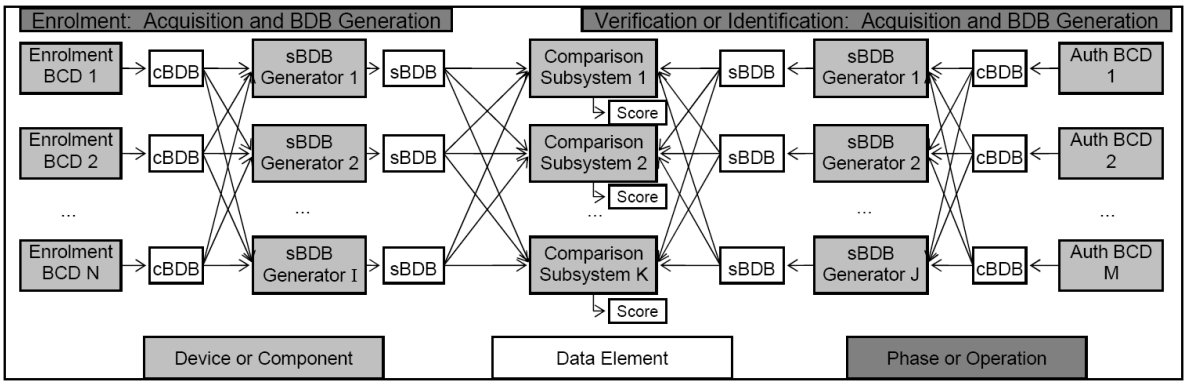

Fig. 1. Testing the performance of interoperable components.

implementation of the PGP message standard[1] is likely to give a deterministic and catastrophic failure, a set of fingerprint minutiae, automatically extracted from a digital image of an analog trait (i.e. the finger), may well give lower accuracy than those those marked by a fingerprint examiner.

This entry gives an overview of the biometric interoperability problem, and introduces the notion that interoperability should properly be quantified in terms of some relevant performance metrics. It proceeds with examples of interoperability challenges which motivates the subsequent contribution on interoperability testing.

Interoperability Challenges

Successful recognition depends on the interoperability of all pieces of equipment used in generating both the enrolled and recognition data records. This begins with the acquisition process, and the primary requirement is that the sensors are interoperable. This typically means that the conversion of the analog human trait into the digital sample produces a defined or commonly understood representation of the original. The following paragraphs give examples that undermine interoperability for the three most common modalities.

FINGERPRINTS: It is common in large scale identity management applications, to acquire flat impressions of a subject's fingers, to associate those with a credential and to verify against one or more of those fingers. While the fingerprint data can be stored in conformance with, for example, the ISO/IEC 19794-4 finger image or 19794-2 finger minutiae standards, subsequent verification attempts depend on the interoperability of the capture devices with the original optical scanner. This is one of few areas of biometrics where sensors are standardized: The U.S. Federal Bureau of Investigation established a physical imaging specification for optical fingerprint sensors. Known as Appendix F [2], this document regulates the imaging capabilities of the acquisition devices such that the representation of the fingerprint ridge structure is accurately represented in the output image, and that the image is defensible in criminal law enforcement. The specification imposes limits on such as the optical resolution of the device the amount of geometric distortion, the imaging area, and spatial uniformity.

FACE: A face image collected at a distance of 30 centimeters is unlikely to be interoperable with another acquired at 1 meter because of the presence of geometric "fish-eye" distortion. This would affect face recognition systems whose internal representation of the face depends on the relative spatial locations of the various anatomical features. While a nonlinear re-sampling of the image could correct such distortion, the resulting spatially varying resolution might undermine accuracy. Another possible solution would be to formulate a mathematical representation that is invariant to this kind of distortion. The actual approach from the commercial and user communities has been to regulate the acquisition process via a formal technical standard. This standard, ISO/IEC 19794-5:2005 *Face Image Data* requires distortion to be absent, and an amendment, ISO/IEC 19794-5/Amd. 1 *Conditions for Taking Photographs for Face Image Data* requires the subject to be positioned at least 0.7m from the camera.

IRIS: The interoperability of three iris cameras was measured in the 2005 ITIRT trial[3]. The results of cross-matching images using a single iris recognition package showed that cross-camera accuracy was generally worse than that for single-camera matching. Possible causes for this, which was not asserted in the report, might be differences in the spectra of the infra-red illuminants, and in the compression applied post capture.

Interoperable data formats

Biometric data interchange standards have been developed to advance interoperability for most of the main biometric modalities. Standards exist for both images and signals, and for "raw" sample data and for processed data. The major extant internationally standardized records are tabulated in Table 1. The standards define a syntactic representation of the data in question. These are usually compact binary encodings of the data suitable for storage on a smartcard or for transmission across a bandwidth-limited communications channel.

Standard	Modality	Processing
19794-2:2005	Fingerprint minutiae	Template
19794-4:2005	Fingerprint	Raw Image
19794-5:2005	Face	Raw or Normalized 2D Image
19794-6:2005	Iris	Raw or Polar Image
19794-7:2007	Signature time series	Multivariate signal
19794-8:2006	Fingerprint skeleton	Processed image
19794-9:2006	Vascular	Raw Image
19794-10:2007	Hand geometry	Binary silhouette Image
19794-13:2010 (est)	Voice	Raw Signal
19794-14:2010 (est)	DNA	Type Signal

Table 1. Biometric data interchange standards for various modalities.

Any interoperability problems that could arise from different implementations of the standards might not be revealed until a test is run or a deployment occurs. While the possible problems are very specific to the standards, the general case is that problems can be expected when two very different sensors are used. For example, in DNA typing, problems would occur if the two sets of loci were disjoint. To avoid such effects the standards variously regulate the biometric acquisition process

		Verification template and matcher provider					
		A	B	C	D	E	G
Provider of enrollment template	A	0.0136	0.0549	0.0458	0.0225	0.0641	0.0417
	B	0.0218	0.0251	0.0385	0.0173	0.0402	0.0192
	C	0.0357	0.0428	0.0225	0.0204	0.0519	0.0348
	D	0.0207	0.0357	0.0301	0.0140	0.0485	0.0316
	E	0.0236	0.0365	0.0340	0.0225	0.0301	0.0286
	G	0.0300	0.0291	0.0447	0.0205	0.0390	0.0129

Table 2. Cross-vendor interoperability for a subset of the minutia detection and matching algorithms evaluated in NIST's MINEX [4] minutia interoperability baseline. The values are false non-match rates at a fixed false match rate of 0.01, for single finger verification on a large offline database.

Interoperability testing

As various interoperability tests were staged around the world [5, 3, 4], the Working Group 5 of ISO/IEC JTC 1's Subcommittee 37 on Biometrics, which standardizes biometric performance tests, started work on interoperability testing. This culminated in 2007, with ISO/IEC 19795-4 *Interoperability Performance Testing*[6] which establishes procedures for the conduct of tests such as those listed.

The standard requires a testing lab to establish and identify one or more application specific figures of merit, such as false non-match and false match rates, and to report them in the manner presented in Table 2. This shows the interoperability of INCITS 378 fingerprint minutia template [7] generators and matchers. It assumes an enrollment template, generated by

equipment identified by the row label, is later verified against a template generated and matched by equipment from the supplier identified in the column headings[1]

The standard also establishes procedures for how a certification body could use mutually low error rates in an interoperability test as a criterion to identify a core group of interoperable products. The standard then addresses how to maintain a certification program in which products are tested in regular ongoing testing campaigns, spanning perhaps several years, in which there might be systematic changes in the difficulty of the test (due to environment, for example).

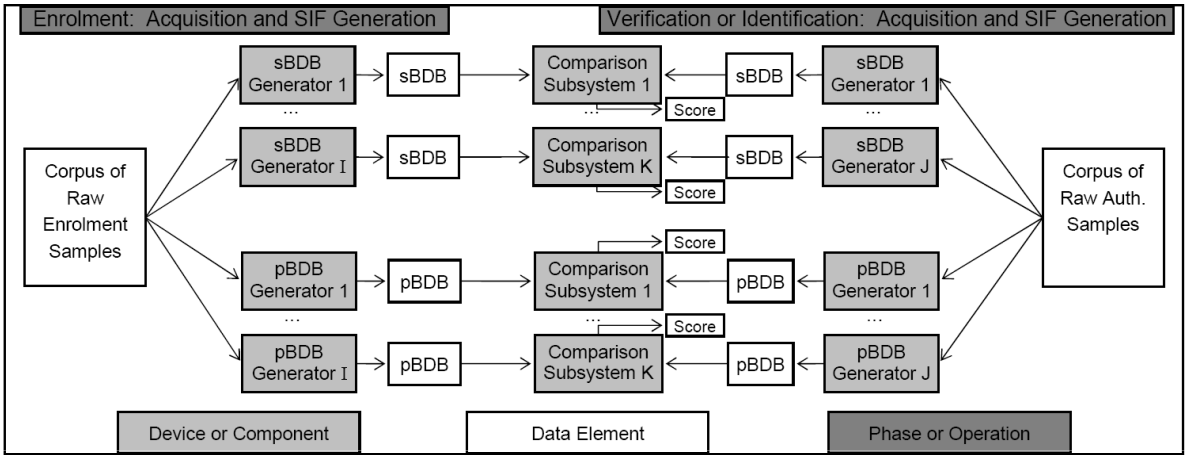

Fig. 2. Testing the sufficiceny of a data interchnage standard. The samples from a fixed corpus are converted to both proprietary and standardized biometric data blocks (pBDBs and sBDBs, repsectively) and then these are recognized by comparison subsystems from the same suppliers.

Sufficiency of a biometric data interchange standard

Biometric data interchange standards support interoperability by allowing developers to implement products producing and processing records conforming to a known format. Such standards define syntactic and semantic representations of the data. For example, a depth value in a 3D face image might be encoded as unsigned integer but the precision might be commercially important; if one supplier can accurately determine depth to within 0.1mm, then they would lose accuracy if the standard only provided for depth resolutions of 0.5mm.

Together the consensus specifications established in a data interchange format standard might offer less accuracy than a totally unconstrained representation of the biometric data, and the quantification of such a loss goes to the *sufficiency* of a biometric data interchange standard[2] by answering the question does standardized data offer accuracies approaching that of unconstrained, non-standard representations.

Figure 2 depicts an offline testing methodology for sufficiency. The MINEX I study[4] quantified sufficiency for the $(x, y, \theta, \text{type})$ encoding of the minutiae defined in INCITS 378[7]. The excerpted results of Table 3 show that proprietary implementations outperform the standard representation by less than the variation between interoperable pairs observed in Table 2.

Summary

While interoperability can be supported by placing appropriate specifications on the various components, particularly sensors, biometric recognition performance tests are sometimes the only means of quantifying interoperability. More importantly, accuracy testing is the most operationally relevant measure of interoperability.

[1] The notable results here are: that the lowest false non-match error rates are lowest when a single company executes all three functions. That this "native mode" gives better performance than the interoperable cases off the diagonal has been attributed to idiosyncratic minutia placement and selection strategies present in minutia detection algorithms.

[2] This terminology is that adopted in the international biometric interoperability performance testing standard[6].

Kind of Template	Provider of template generator + matching algorithm				
	A	B	D	E	G
Proprietary	0.0089	0.0189	0.0089	0.0251	0.0047
Standard	0.0136	0.0251	0.0140	0.0301	0.0129

Table 3. False non-match rates at fixed a false match rate of 0.01 for fingerprint verification algorithms using fully proprietary, and formally standardized, templates. The proprietary accuracies would only be available in an interoperable application if the images were exchanged.

Related Entries

Performance Testing, Standards, Interoperability

References

1. J. Callas et al.: RFC 4880 - OpenPGP Message Format. (2007) http://tools.ietf.org/html/rfc4880.
2. T. Hopper et al.: IAFIS Image Quality Specifications, EBTS Appendix F. Technical report, FFBI Criminal Justice Information Services Division (2008) http://www.fbibiospecs.org.
3. M. Thieme et al.: Independent testing of iris recognition technology final report. Technical report, International Biometric Group (2005) http://www.biometricgroup.com/ITIRT/.
4. P. J. Grother et al.: Minutiae Exchange Interoperability Test MINEX - Performance and Interoperability of the INCITS 378 Fingerprint Template. Technical Report NIST 7296, National Institute of Standards and Technology, Gaithersburg, Maryland (2006) Available at http://fingerprint.nist.gov/minex.
5. J. Campbell et al.: Seafarers' Identity Documents - Biometric Testing Campaign Report. Technical report, International Labour Organization (2005)
6. JTC 1, SC37 Biometrics, Working Group 5: ISO/IEC 19795-4 Interoperability performance testing. (2007) http://webstore.ansi.org.
7. INCITS M1, Biometrics: INCITS 378:2004 Fingerprint minutia format. 1 edn. (2004) American National Standard for Information Technology.

Definitional Entries

Biometric Data Interchange Standard

A published documentary specification of a data record for clear exchange of subject's biometric data between two parties.