

1. Background

Quality measurement plays vital role in improving biometric system accuracy and efficiency during the capture process (as a control-loop variable to initiate reacquisition), in database maintenance (sample update), in enterprise-wide quality assurance surveying, and in invocation of quality-directed processing of samples. If quality can be improved, either by sensor design, by user interface design, or by standards compliance, better performance can be realized. For those aspects of quality that cannot be designed-in, an ability to analyze the quality of a live sample is needed. This is useful primarily in initiating the reacquisition from a user, but also for the real-time selection of the best sample, and the selective invocation of different processing methods.

Biometric quality analysis is a technical challenge because it is most helpful when the quality measures reflect the performance sensitivities of one or more target biometric matchers. NIST addressed this problem in August 2004 when it issued NIST Fingerprint Image Quality (NFIQ) algorithm. NFIQ is a fingerprint quality measurement tool; it is implemented as open-source software; and is used today in U.S. government and commercial deployments. NFIQ is a machine-learning algorithm and its key innovation is to produce a quality value from a fingerprint image that is directly predictive of expected matching performance. NFIQ serves as a publicly available reference implementation. With advances in fingerprint technology since 2004, an update to NFIQ is needed.

2. Statement of Need

NFIQ has been designed to be matcher independent. For applications where matching algorithm is not known or subject to change, a generalized (i.e. matcher independent) quality is needed. However, when comparison algorithm is known use of quality assessment algorithm tuned to predict the performance of the comparison algorithm is more suitable. Therefore the next generation of finger image quality, should provide both options of “generalized” (i.e. matcher-independent) or “dedicated” (i.e. matcher dependent).

Interoperability of quality scores is another challenge in exchange of quality scores. Part 1 of the multipart ISO/IEC 29794 Biometric sample quality standard defines a binary record structure for the storage of a sample’s quality data. It establishes requirements on the syntax and semantic content of the structure. Specifically, it states that the purpose of assigning a quality score to a biometric sample shall be to indicate the expected utility of that sample in an automated comparison environment. That is, a quality algorithm should produce quality scores that target application-specific performance variables. For verification, the metric would usually be false-match and false-non-match rates that are likely to be realized when the sample is matched. This, by itself, is not sufficient for accurate interpretation of quality scores generated by different quality assessment algorithms and some normalization or calibration is needed.

The goal of this workshop is to address the technical short-falls in current quality assessment technology, and aims to engage industry to improve core finger image quality assessment technology based on lessons learned from recent deployments of quality assessment algorithms (including NFIQ) in large-scale identity management applications.

3. Finding a way forward

We see three possible options for future of NFIQ.

1. The first and probably least attractive option is to do nothing. Stop the work on NFIQ.
2. The second and more reasonable option is to do incremental upgrade/revision to NFIQ 1.0. Examples of such improvements are: inclusion of core in the feature vector (to overcome the issue with impressions of tip only); more resolution (0-100 instead of 1-5 levels); re-train with newer data and/or newer matcher; and train on rolled impression.
3. The third, and our proposed option is, in consultation and collaboration with industry, implement various modular versions of NFIQ. This option is explored in more details below.

3.1. Proposed scientific approach for modular NFIQ

We propose development of a set of standardized finger image quality components plus multivariate statistics techniques to relate biometric performance metrics such as false non-match to the standardized or vendors' generated quality components (i.e. features). The outcome will be modular quality algorithms (NFIQ 2.0) with plug-and-play feature vector extraction algorithms that could be performed by a vendor, an open source algorithm, or a combination of the two.

This task requires close collaboration with providers.

Development of modular NFIQ 2.0, which by design, could be either generalized or specialized to a particular comparison algorithm. As with NFIQ, the new NFIQ 2.0 will have two major computation steps:

- Feature extraction, and
- Training of a machine-learning algorithm.

Feature extraction consists of measuring appropriate image characteristics that convey information for comparison algorithms. The feature set is comprised of elements such as local noise, continuity of ridge flow, area of the finger image impression, and number of minutiae. A feature vector is computed from each image and its component are combined using a trained machine-learning algorithm so that the image quality score is reflective of positive or negative contribution of the sample to the overall performance of the system.

Modular NFIQ allows for dedicated quality assessment (feature extraction and training tuned to a particular algorithm) or generalized and matcher independent quality assessment to be used in any application without the knowledge of the deployed (template generation or matcher) deployed. A brief description of our proposed NFIQ 2.0 straw-man design follows.

3.1.1. Feature Extraction

Modular NFIQ 2.0 supports plug-and-play feature vectors. Three separate options are considered here:

1. Black Box Feature Vector

Vendors can supply their proprietary (black box) feature extraction software as an .exe or a compiled library. This converts an image into a proprietary quality vector. This option generates a "dedicated" (i.e. extraction algorithm dependent) quality assessment tool. Incorporating vendor proprietary feature vectors, improves the efficiency of quality computations in the field because the feature vector computation is potentially a by-product of template generation. Vendors' intellectual property is reserved. The success of black box plan depends on vendor participation.

2. Clear Box Feature Vector

Optionally, in collaboration with industry, a set of quality components will be defined and perhaps formally standardized. These quality components shall model failure modes and sensitivities of current fingerprint matching algorithms. Examples include: zonal quality, clarity of ridges, size of fingerprint, or number of minutiae. NIST will develop open-source reference implementations for standardized quality components.

3. Gray Box Feature Vector

This option consists of a subset of standardized feature vector (clear box) appended to a vendor-supplied proprietary feature vector (black box). Vendors can opt for some of the standardized quality components, and extend that with their own proprietary quality components.

3.1.2. Training a machine learning algorithm

Training looks for structure in the data and ultimately building a model to relate the response variable (e.g. performance as false non-match rate, or area-under-ROC-curve) to the exploratory variables (i.e. features or quality components). We explore different multivariate statistical techniques to obtain the optimal model. Training could be customized to a comparison algorithm or generalized to a class of comparison algorithms.

Training will be performed by NIST. NIST will return the detail information on the machine-learning algorithm including trained weights to the vendor. NIST will use some of its diverse set of finger images for training. Avoiding over-fitting problem and building models that could explain structures of unseen data is an important and difficult problem in training machine-learning algorithms. Use of a large and diverse set for training can mitigate this problem. NIST will calibrate quality scores either as part of training or afterwards, and therefore ensuring interoperability of final quality scores.

The outcome will be a family of quality algorithms that could be application-independent or tuned to particular applications. Interoperability is achieved by uniform interpretation of quality scores; therefore, it expands a marketplace of interoperable products.

Figure 1 summarizes our proposed approach to NFIQ 2.0.

3.2. Benefits

The advantages of incorporating vendors' provided features are a) improved efficiency (lower computation time) and b) better tuning of a quality algorithm to a specific application.

The advantages of NIST developing multivariate statistical techniques to linearly or non-linearly combine the features into a quality score are a) ensuring interoperability of quality scores, and b) use of NIST's large and diverse data will result in development of more robust models and avoid the problem of over-fitting to small data.

NIST will calibrate quality scores either as part of training or afterwards, and therefore ensuring interoperability of final quality scores.

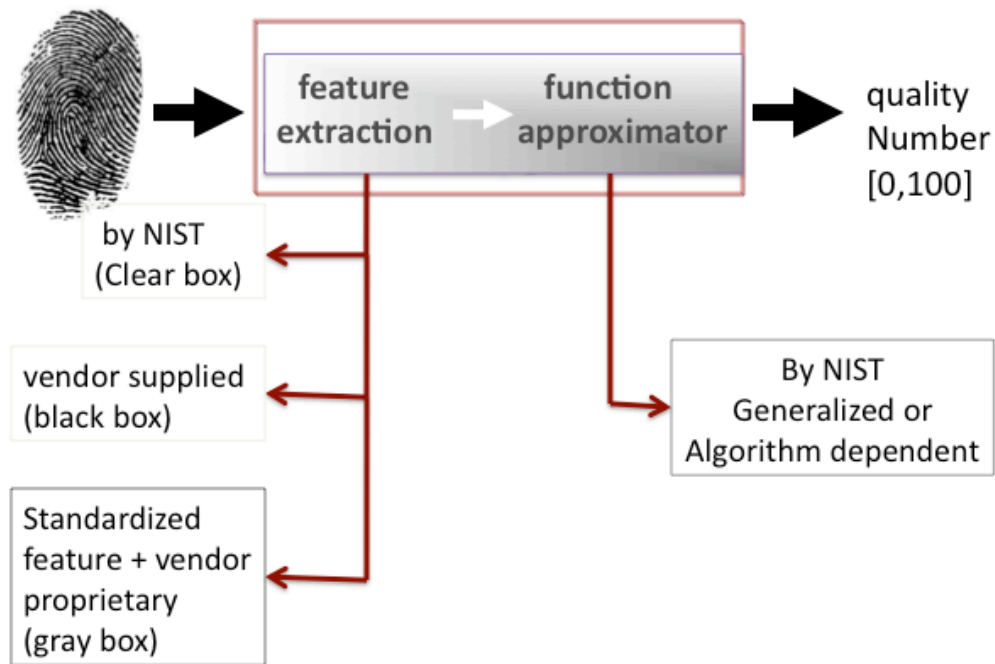


Figure 1. NIST proposed NFIQ 2.0