Monitoring and Reporting of Fingerprint Image Quality and Match Accuracy for a Large User Application

Teddy Ko and Rama Krishnan
Lockheed Martin
1120 Vermont Avenue, NW – 5th Floor
Washington, DC 20005
{*teddy.ko, rama.krishnan*}*@lmco.com*

Abstract

The main objective of this paper is to present the methodology used for measuring and monitoring the quality of fingerprint database and fingerprint match performance of a large user fingerprint identification system. The Department of Homeland Security's (DHS) Biometric Identification System is used as an example for this study. In addition the paper will present lessons learned during system performance testing and independent validation & verification analysis of large scale systems such as DHS's Biometric System and recommend improvements for the current test methodology.

1. Introduction

An automated fingerprint identification system traditionally consists of three subsystems: fingerprint image acquisition, fingerprint feature extraction, and fingerprint matching. In image acquisition, a digital image of a fingerprint is captured either from the live scan of a person's finger or from an inked impression of a person's fingerprint on paper (e.g., fingerprint cards). Feature extraction process captures the unique characteristics (referred to as minutiae) in some condensed form to facilitate matching. Matching involves computing the likelihood of the fingerprint coming from the subject with persons enrolled (stored) in the database. The performance of the whole system depends on how well each subsystem behaves.

In a live, operational fingerprint identification system, the feature extraction and matching subsystems of a specific vendor are used and this forms a fixed component. The fingerprint acquisition subsystem, which is driven by fingerprint capture interface implemented in the capture workstation, is a variable component, and can impact quality. Recent studies have demonstrated that all fingerprint identification systems tested have widely varying matching performance, which depends, among other factors, on fingerprint image quality [3]. The captured fingerprint quality has a direct effect on matcher accuracy. Therefore operational systems need effective image quality metrics for real-time operator feedback, to ensure the capture of acceptable quality images to maintain high identification accuracy [1].

A perfectly scanned or sensed fingerprint image under ideal conditions has clear minutiae (ridges and valleys) used for the fingerprint matching process. An automated fingerprint identification system can perform well on such fingerprint images. However, often skin condition or imperfect acquisitions cause the captured fingerprint image to be far from ideal quality. Unclean sensor plates, non-uniform and inconsistent contact can result in the capture of poor quality images that will have a negative impact on identification performance accuracy.

There are two parameters: false reject rate (type 1 error rate) and false accept rate (type 2 error rate) that are used to measure the performance of a biometric identification system. For any large user biometric identification application, it is possible to measure the false accept rates of the operational system; but is hard to measure the false reject rates in operational system as these cannot be detected in the system. The false reject error rates of an identification system are typically established in a formal but off-line benchmark test of the system. With the emergence of large scale identification systems such as IDENT/US-VISIT system, it becomes critical to have a theoretically and empirically sound methodology for measuring and estimating the match performance characteristics of a large user identification system. This helps to ensure that the deployed systems are operating at or close to their optimum performance levels.

It is also desirable to assess the quality of a fingerprint image in real time as a quality control

Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop (AIPR'04)
1550-5219/04 $ 20.00 **IEEE**

procedure. This allows poor image acquisition to be corrected through recapture and facilitates the capture of best possible image within the capture time window configured in the system. This results in the capture of more good quality fingerprint images that can improve the system identification accuracy and the integrity of fingerprint database.

The paper will describe the methodologies used for estimating the identification accuracy of large-scale identification system through empirical means and comparison of these measurements with the actual performance observed in the operational system and an offline benchmark test. This will help to ascertain if the performance characteristics of the operational system are better or at least consistent with expected performance figures for the system. The paper will also describe a fingerprint enrollment update process – best minutiae update – used to replace the currently enrolled fingerprints of a subject with better quality fingerprints of the same subject captured subsequently in the system. This update process has the desirable effect of improving the overall quality of the fingerprint database over time and thereby improving the overall system match accuracy.

1.1. Objective

The main objective of this paper is to present the methodology used for measuring and monitoring the quality of fingerprint database and fingerprint match performance of a large user fingerprint identification system. The Department of Homeland Security's (DHS) Biometric Identification System – IDENT/US-VISIT – will be used as an example for this study. In addition, the paper will also present some recommended procedures relevant from image quality assurance perspective.

2. Fingerprint Quality Assessment

A fingerprint is a pattern of friction ridges on the surface of a fingertip. A good quality fingerprint has distinguishable pattern and features that allow the extraction of features that are useful for subsequent matching of fingerprint pairs. It is known that most fingerprint matcher algorithms in use are sensitive to clarity of ridges and valleys, measures of number and quality of minutiae, and size of the image [1].

Poor fingerprint image quality is a problem when it contributes to the inability of a machine or human expert to identify minutiae points or core/delta points in a given fingerprint image. Such failure to extract minutiae points is usually caused by poor ridge (valley) flow, poor contrast and brightness in the image, or small, partial images.

Most of currently available fingerprint image quality assessment or analysis algorithms have been designed to mimic the human visual perception of fingerprint image quality. The algorithms assess the percentage or size of the given image that may contain actual fingerprint and how reliable the ridge flow could be detected from the located fingerprint area. Although most of algorithms use similar attributes in determining the quality of the fingerprint images, but the scoring functions between algorithms are often different – some tailored for the benefit of machine matching and some for human inspection.

It has been established that fingerprint match performance of a system is directly dependent on the quality of the fingerprint images used for the match. Therefore, it is critical that fingerprint capture process includes real time quality assessment and feedback to improve the quality of fingerprints captured in the system to maintain high identification accuracy. Brief descriptions of two fingerprint image quality assessment algorithms are presented in the next section for better understanding of the principles used for image quality assessment.

2.1. NIST Fingerprint Quality Assessment Algorithm

The minutiae detection (MINDTCT) package of NIST Fingerprint Image Software (NFIS) [2] has a fingerprint minutia detector algorithm that accepts a fingerprint image and automatically detects minutia. It also assesses minutia quality and generates an image quality map. To locally analyze a fingerprint image, NFIS divides the image into grid of blocks. To access the quality of each block, NFIS computes several maps (direction map, low contrast, low flow, and high curve) and summarizes the result in a quality map.

The purpose of direction map is to determine the areas of the image with sufficient ridge structure. Well-formed and clearly visible ridges are essential to reliable detection of ridge endings and bifurcations. The low contrast map is used to flag where the blocks have sufficiently low contrast. The low contrast map separates the background of the image from the actual fingerprint areas, and maps out smudges and lightly inked areas of the fingerprint. Minutiae are not detected within low contrast blocks in the image. Low flow map marks the blocks that could not initially be assigned a dominant ridge flow. Minutiae detected in low flow areas are not reliable. A high curve map is used to mark blocks that are in high curvature areas of the

fingerprint. Minutiae detected in high curvature areas are not reliable, and this is especially true of the core and delta regions of a fingerprint. The low contrast map, low flow map, and high curve map all point to different low quality regions of the image. The information in these maps is integrated into one quality map, and contains 5 levels of quality (4 being the highest quality and 0 being the lowest).

For each fingerprint image, the MINDTCT generates its quality map. Blocks with quality 0 are regarded as background. The blocks with quality 1 or better are considered the effective size of the fingerprint image or foreground. The percentage of foreground blocks with quality 1, 2, 3, and 4 are computed. Fingerprint images with higher number of quality zone 4 (equivalently smaller number of quality zone 1 and/or 2) are more desirable.

2.2. Cogent Fingerprint Quality Assessment Algorithm

The matchers used in the current IDENT/US-VISIT system are developed and commercially marketed by Cogent Systems. Cogent also provides software for the calculation of fingerprint image quality. The Cogent image quality measure is based on a scale from 1 to 8, where 1 is the best image quality value and 8 is the worst image quality value. These quality values are grouped in three quality classes in the operational IDENT/US-VISIT system: 1 to 4 as good quality, 5 to 6 as average quality, and 7 to 8 as poor quality.

NIST studies have found Cogent image quality to be a good rank statistic for all the algorithms and all data set used in NIST's test [3]. The error rate of the best (quality 1) fingerprints is always lower than the error rate of any other image quality level, and the error rate of the worst (quality 8) fingerprint is always the highest. All other image quality levels result in the expected level of the error rates.

2.3. Fingerprint Image Quality and Match Accuracy Correlation

In an Automated Fingerprint Identification System (AFIS), the input can be a single fingerprint, two index fingerprints, or up to ten fingerprints and the output is a list of potential match candidate subjects, whose fingerprint were found to be similar to the searched fingerprint inputs based on the fingerprint match scores.

In biometric matching studies, the performance of the system is expressed by the accuracy of the system. In a biometric decision, a type of Yes/No pattern recognition decisions, there are four possible outcomes: True Accept (**TA**) or called correct accept, False Accept (**FA**), False Reject (**FR**), and True Reject (**TR**) or called correct reject. FA and FR are errors, while TA and TR are correct outcomes sought in a biometric system. False Accept Rate (**FAR**) and False Reject Rate (**FRR**) are widely used standard metrics of the identification/verification accuracy of biometric systems. The performance of a biometric system are usually shown as a Receiver Operating Characteristic (ROC) curve that plots the true accept rate vs. false accept rate at different match score thresholds. By manipulating the decision criteria, the relative probabilities of these four outcomes can be adjusted in a way that reflects their associated costs and benefits to a biometric identification system. Reliability and uniqueness of features are two dominant parameters that contribute to FARs and FRRs in automated fingerprint verification/identification.

An independent and formal test conducted by the NIST using data gathered from the IDENT/US-VISIT system and our independent verification & validation (IV&V) analysis of operational data have both shown that the better captured fingerprint image quality will have better match accuracy. The image quality score, e.g., Cogent fingerprint image quality score, is a good predictor for the match performance.

3. Approach

To accurately measure the match accuracy of a large user application of automated fingerprint identification system, large number of probes need to be searched against a large background gallery size in the millions. The results [3] of experiments reported in Matching Performance for the IDENT/US-VISIT System Using Flat Fingerprint used a probe set of 60,000 fingerprint pairs searched against a background gallery of 6 millions. This resulted in a test that required hundreds of billion raw matches when using two index fingers, and it took days of CPU time to complete. This demonstrates that special computer hardware matchers and special software implementation are essential and needed for conducting large scale AFIS testing. It is important for identification application with a large user population to use same or similar test protocol that NIST has established to measure the system's match accuracy from time to time to ensure the baseline performance has been improved or at least maintained. However, due to the above protocol requires heavy resources and computation time, other alternatives are needed for real-time or frequently needed match performance monitoring and reporting.

In operation, we have derived and designed methods and procedures to measure and estimate the system's matching performance based on some of conclusions that NIST has drawn from the results of the formal test it has conducted on the existing IDENT/US-VISIT one-to-many and one-to-one matching systems. In the one-to-many matching, the false accept rate (FAR) using index finger pairs is very close to linearly increasing with the database size. In both one-to-one and one-to-many matching, the Cogent image quality measure is a good predictor of matching performance.

Our technical approach analyses the match results of both one-to-one and one-to-many search transactions recorded in the operational database and measures the identification performance characteristics observed in the operational system. The quality of the fingerprint images associated with all the operational match transactions are analyzed to determine image quality-match performance correlation in the operational systems. The operational performance characteristics can now be compared with estimated baseline system performance characteristics to measure the correlation and consistency between the two measurements.

4. Fingerprint Image Quality Measurement

For a large user fingerprint identification system, the image quality measurements can be used not only to assess the system's match performance but also to identify the sources experiencing quality control problem leading to poor quality fingerprint image submissions. Remedial actions can be instituted at these sites to address the image quality issue.

4.1. Image Quality by Application

The IDENT/US-VISIT identification system supports multiple applications, like Visit Entry, Visit Exit, and Department of State's Biometric Visa, etc. Each application has its own selected fingerprint scanners, capture software application, and different environment configurations and settings. All of these differences will have impacts on the overall quality of captured fingerprint images. Thus, it is important to monitor and report the image quality distribution according to different application. This helps to determine if there is an image quality problem specific to an application for possible corrective action.

4.2. Image Quality by Site/Terminal

There can be variation in image quality from the different site/terminal within the same application resulted from operator training, site configuration and operational condition of the fingerprint scanners used at the site. The reporting of image quality distribution by site/terminal identifies abnormal terminals or sites for further investigation and corrective action.

4.3. Image Quality by Capture Device

It has been shown [4] that due to the use of different scanner physics principle, mechanical design, GUI capture interface, etc., there can be variation in the quality of captured fingerprint images between capture devices. The reporting of image quality distribution by scanner type identifies any scanner specific image quality issues.

If any capture device at a specific site shows worse quality capture statistics compared to average quality statistics of the system, further investigation should be initiated at the specific site for corrective action. If the image quality capture statistics does not improve after the adjustments and/or corrections, the specific scanner should be discontinued from use to eliminate the degradation in overall of system image quality.

4.4. Image Quality by New or Repeated Subject

It is a well known fact that once the users get more familiar with the interaction with the fingerprint input device, their fingerprint presentation will get better resulted in improved fingerprint quality. The results of the weekly US-VISIT image quality distribution reports show improved quality fingerprint capture statistics from repeat visitors, compared to the first time visitors. This data shows that enhanced fingerprint capture training manuals with best quality fingerprint capture tips might alleviate the "first time user" syndrome.

4.5. Image Quality by Matcher Enrollment

The matching performance is affected by both quality of fingerprint images submitted to the system for search and the quality of fingerprint images already enrolled in the databases of the matchers used for matching. A capability to show the weekly/monthly image quality distribution snapshots of the matcher image database is important. The image quality distribution snapshots of the matcher will show if the image quality is improving or getting worse or remaining consistent over time, and this information

will be useful in ascertaining the expected match performance over time.

The matcher fingerprint image quality snapshot report captures the monthly matcher fingerprint image quality statistics, and shows the monthly match image quality improvement trend resulting from the best minutiae update process.

4.6. Image Quality by Finger and between Fingers

In this report, we examine the image quality distributions of right and left fingers, and the image quality correlation of the right and left fingers. The results of this report can be used to determine if we need to adjust the system's single finger match score thresholds differently for right or left finger to compensate for image quality correlation differences for the two fingers. The differences in image quality distribution for the two fingers could also result from the scanner location ergonomics at the capture sites, which may be more conducive for the presentation of the left or right finger. If this is indeed the cause for the variation in image quality capture statistics for the two fingers, corrective action can be initiated to rectify this finger ergonomics problem.

4.7. Image Quality Trend Analysis

The image quality trend analysis report provides the image quality capture statistics of all sites/terminals over time. The trend analysis reports can be used to identify to any downward trends in image quality or sudden changes at any sites for further investigation.

5. Match Accuracy Measurement

In biometric matching studies, the performance of the system is expressed by the accuracy of the system. Verification is defined as a one-to-one match used to decide if the individual is who he/she claims to be. Identification is defined as a one-to-many match designed to determine if a specified subject is in the database. False Accept Rate (**FAR**) and False Reject Rate (**FRR**) are widely used standard metrics of the match accuracy of biometric systems.

In an operational fingerprint identification system, we will not be able to directly measure the False Reject Rate (FRR), as missed identifications cannot be detected in the system. However, studies carried out by NIST and our independent verification and validation (IV&V) tests on operational IDENT and US-Visit systems have demonstrated that we can use the fingerprint image quality to estimate, predict the fingerprint match performance characteristics.

5.1. Verification One-to-One Match Analysis

The purpose of the verification (1:1 match) report is to provide the verification fingerprint match performance summary for the operational system for the specific time period. The report provides the verification false reject rate (FRR) observed in the operational system. The report may be provided as a weekly, monthly, or as an ad hoc as needed report.

To compute the verification FRR for a specific time period from the operational system, we need to find out the counts for the total verification transactions (vTT), transactions with confirmed positive matches (vTPM), and verification transactions with no match (vTNM) for the specific time period. The verification false reject rate (vFRR) = (vTNM/vTT)*100%, where vTT = (vTPM +vTNM).

In addition to the verification FRR, the report also presents the results of the detailed analysis of image quality match correlation. The verification match related data including (1) image quality of the verification subjects and match candidates, (2) match scores and image quality scores correlation. In order to facilitate this analysis, all the relevant data associated with the verification match transactions are extracted from the operational database for detailed analysis.

5.2. Identification (Open Search) One-to-Many Match Analysis

The purpose of the open search (1:N match) report is to provide the open search fingerprint match performance summary for the operational system for the specific time period. The report provides the search false accept rates (FARs) observed in the operational system. The report may be provided as a weekly, monthly, or as an ad hoc as needed report.

To compute the search FAR for a specific time period from the operational system, we need to find out the counts for the total search transactions (sTT), search transactions with positive matches (sTPM), search transactions with confirmed false matches (sTFM), and search transactions with no match (sTNM) for the specific time period. The search false accept rate (sFAR) = (sTFM/sTT)*100%, where sTT = (sTPM+sTFM+sTNM). Confirmed false match is any match that is designated as false match by the candidate verification fingerprint examiner.

In addition to the search FAR, the report also presents the results of the detailed analysis of open

search related data including (1) image quality of the search subjects and match candidates, (2) match scores and image quality scores correlation. In order to facilitate this analysis, all the relevant data associated with the open search transactions are extracted from the operational database for detailed analysis.

6. Lessons Learned and Recommendations

The matching performance characteristics of an identification system are directly affected by the quality of fingerprint images captured and present in the database. Different type of fingerprint capture methods and devices result in the capture of fingerprint images with varying quality. Variations are also possible in the images of a same subject captured from different acquisitions from a same fingerprint scanner. Image quality assurance is a challenging and critical task to ensure the capture of good quality fingerprint in the system to ensure high identification performance. Thus, an evaluation test methodology to determine if the new scanner planned for integration into the system meets the performance criteria required for successful interoperability with the existing system needs to include statistically valid test procedures [4].

The factors that can affect fingerprint image quality include the fingerprint scanner used, application's fingerprint capture GUI implementation, environment, subject age, and cooperative or non-cooperative user application. All these factors needed to be considered in fingerprint image quality studies.

6.1. Fingerprint Scanner Interoperability

Recent studies on biometric accuracy requirements for large scale identification application – conducted by accredited independent testing agencies, including NIST – have addressed the impact of fingerprint image quality on both verification and identification accuracies. Therefore it is critical to ensure that the fingerprint images captured by a new fingerprint scanner that is integrated into an existing identification application are interoperable with the images captured from the existing fingerprint scanner devices and result in similar (if not better) identification performance characteristics.

6.2. Best Minutiae Update

Best minutiae update is a fingerprint enrollment update process which replaces the currently enrolled fingerprints of a subject with better quality fingerprints of the same subject captured subsequently in the system. This process has the desirable effect of improving the overall quality of the fingerprint database over time and thereby improving the overall system match accuracy.

7. Conclusions

This paper has presented the methodology used for measuring and reporting quality of fingerprint database and fingerprint match performance of a live large user fingerprint identification system. The proposed methodology uses some of the concepts discussed in NIST's formal test reports on the IDENT/US-VISIT system.

The test methodology describes the approaches for measuring false reject rate and false accept rates in a live large user identification system.

The proposed fingerprint image quality monitoring and reporting methodology addresses different aspects related to image quality. The reporting methodology not only provides overall fingerprint capture statistics by application/site but additional trend analysis information to identify abnormal or downward trends in image quality in the systems for timely corrective action. These measures will facilitate the maintenance of overall fingerprint quality at acceptable levels to achieve high identification accuracy for the system.

References

[1] E. Tabassi, C.L. Wilson, and C.I. Watson, "Fingerprint Image Quality", NIST technical report NISTIR 7151, August 2004.

[2] M.D. Garris, C.I. Watson, R.M. McCabe, and C.L. Wilson, "User's Guide to NIST Fingerprint Image Software (NFIS)", NIST technical report NISTIR 6813.

[3] C.L. Wilson, M.D. Garris, and C.I. Watson, "Matching Performance for the US-VISIT IDENT System Using Flat Fingerprints", NIST technical report NISTIR 7110, May 2004.

[4] T. Ko and R. Krishnan, "A Look Beyond Data Exchange to Fingerprint Scanner Interoperability", Biometric Consortium Conference, Washington DC, September 2003.

[5] T. Ko and R. Krishnan, "Fingerprint and Face Identification for Large User Population", *Journal of Systemics, Cybernetics and Informatics*, Vol. 1, No. 3, 2003, pp. 87-92.

[6] K.W. Bowyer and P.J. Phillips, *Empirical Evaluation Techniques in Computer Vision,* IEEE Computer Society Press, 1998.