# Analyzing Fingerprints of Indian Population Using Image Quality: A UIDAI Case Study

M. Vatsa[1], R. Singh[1], S. Bharadwaj[1], H. Bhatt[1], R. Mashruwala[2]
*1 - IIIT Delhi, India   2 - UIDAI, India*

## Abstract

*In large scale deployment of fingerprint recognition systems, especially in Indian environment, there are some challenges involved. Along with sensor noise and poor image quality, presence of scars, warts, and deteriorating ridge/minutiae patterns in fingerprints from rural population affect the data distribution. In other words, quality of fingerprint patterns, particularly belonging to rural Indian population, may differ from standard urban or western population and may be difficult to process. Since there is no study that analyzes fingerprint images in Indian context, this paper presents an analytical study using standard fingerprint image quality assessment tool and fingerprint databases collected from the rural and urban Indian population. On a database of over 0.25 million images, we observe the patterns that are worn and damaged cause poor quality ridges and therefore can affect the performance. Also, region specific causes such as manual labor and Lawsonia Inermis also degrade the quality of fingerprints.*

## 1. Introduction

Automatic Fingerprint Identification Systems (AFIS) are used in several large scale applications across the world for identifying individuals [1]. In real time applications, AFIS should be stable and robust to environmental dynamics and variations in data distribution. However, sensor noise, poor quality, and partial capture may affect the performance. Moreover, success of AFIS is dependent on the interaction between an individual and sensor and it can cause variations in image quality and biometric features. Therefore, in general, we can assert that poor quality images contribute to difficulty in detecting features from the image and hence may decrease the recognition performance.

In Indian environment, fingerprints/biometrics has not been used for any large scale civilian applications. Recently, Indian Government has set up Unique Identification Authority of India (UIDAI) to issue Unique Identification (UID) numbers to all residents in the country [3]. In this very large scale civilian project, biometrics specifically fingerprint recognition is one of the major components. Before it is implemented, there is a need to analyze and assess feasibility of using different biometric modalities. We believe that Indian environment is quite different compared to western and European countries. Indian conditions are unique in two ways:

- Larger percentage of population is employed in manual labor, which normally produces poorer biometric samples.
- Biometric capture process in rural and mobile environment is less controllable compared to the environmental conditions in which Western data is collected.

There is no study that (1) estimates the accuracy achievable for fingerprints under Indian conditions or (2) analyzes the quality of fingerprint images and provides insight about *data quality*. This paper focus on analyzing fingerprints using image quality tools for UID application. Specifically, using the fingerprint images collected from Indian population, we computed NFIQ quality scores [2] of images and analyzed the quality distribution of Indian fingerprints.

## 2 Databases and Protocol

We focused on the ability to leverage image quality assessment tools in (1) analyzing the input biometric samples that are obtained from diverse, disparate sensors and (2) characterizing the samples based on the quality and amount of information present. Using three fingerprint databases and NFIQ tool, fingerprint image quality based experimental evaluation was performed.

1. DB1: Prepared by the authors, contains images from 27 urban individuals (total 27*10*5 = 1350 images) and 81 rural individuals (total 81*2*10 = 1620 images). This database is prepared using single impression sensor.

2. DB2: This database contains real world slap images from over 20,000 individuals. Each slap fingerprint image was segmented using commercial fingerprint segmentation tool. After segmentation, this database contains over 0.2 million unique fingerprint images.

3. DB3: Another database with real world fingerprint images. This database contains images from rural India pertaining to 5600 individuals (around 56,000 unique images) and are segmented using commercial software.

Note that DB2 and DB3 databases contain only single impression per finger[1]. Using these database, NFIQ image quality based experimental analysis was performed. Specifically, the experiments were divided into four steps:

**Step 1:** DB1 is divided into eight bins according to gender, age group and urban/rural whereas in DB2 and DB3, four groups were created according to gender and age, i.e.

- Group 1: Male, 18 to 40 years age
- Group 2: Female, 18 to 40 years age
- Group 3: Male, 40+ age
- Group 4: Female, 40+ age

**Step 2:** Using NFIQ image quality assessment tool, each finger in the different bins is assigned a quality score. NFIQ quality score of 1 represents good quality image and 5 represents poor quality.

**Step 3:** For each group, quality score distribution is computed. We also computed the score distribution for each finger. This provides an insight whether all fingerprints are good for recognition purposes or only some selected ones. Finally, we computed the score distribution for the complete database.

**Step 4:** We visually inspected images and distribution models to get insight into poor quality images. Specifically, we (1) analyzed whether the poor image quality is due to irregularities in the capture process and is it correctable, (2) analyzed quality variation across fingers, and (3) computed correlation of quality score across multiple fingers of the same person.

[1] Most of the existing fingerprint databases that are caputred in rural settings and prepared by government agencies, do not have multiple samples of fingerprints.

## 3. Experimental Analysis

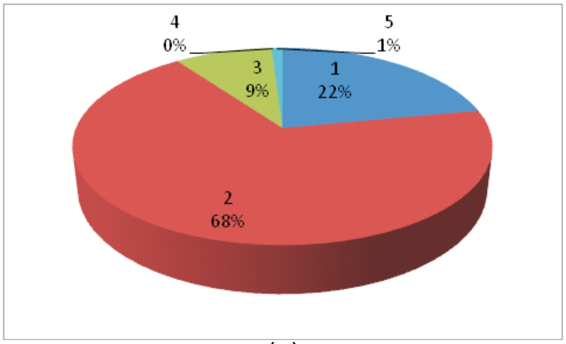
Figure 1 shows the NFIQ quality score distribution on the three databases. These NFIQ based experiments provide an insight of fingerprint quality in the Indian environment. The key results and analysis of our experiments are summarized below.

1. NFIQ results on the databases are encouraging especially if the fingerprint images are captured under controlled environment. For the majority of images, quality scores are good with NFIQ score range of 1-3. Usually with fingerprint images of NFIQ score 1-3, fingerprint feature extraction and matching algorithms provide high accuracy. Therefore, to achieve good recognition accuracy, it is desired to collect good quality images using optimized operational mechanism and good sensors.

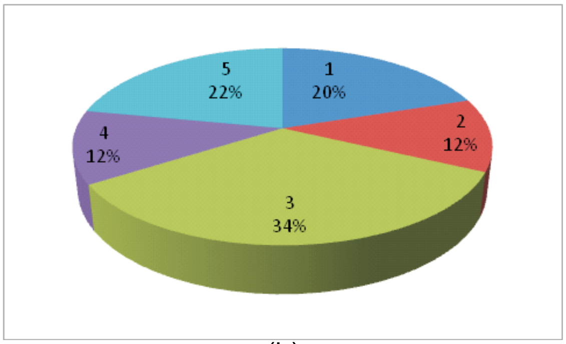
2. The commercial segmentation tool failed for some fingerprint images. Especially, as shown in Figure 2, due to *ghost* fingerprint, segmentation tool fails to correctly segment the fingerprint images. It is also observed that the slap fingerprint segmentation tools require some prior training on the Indian databases. After training, the segmentation results improve by 3%. This also suggests that in deploying a biometrics (fingerprint) system, a carefully designed a priori training dataset and procedures will help in improving the performance.

3. Since the NFIQ tool is trained using western data, we observed that there are around 4.6% errors in correctly assessing the quality scores. For example, as shown in Figure 3 some poor quality images are assigned a score of 1. If such images are used during enrollment, then verification (typically will be performed with single finger devices) will be very poor because of the low area of overlap. Further, if this person applies for another UID and the second time the fingerprints are correctly captured, there will be low overlap with the first capture and hence the duplicates will not be found.

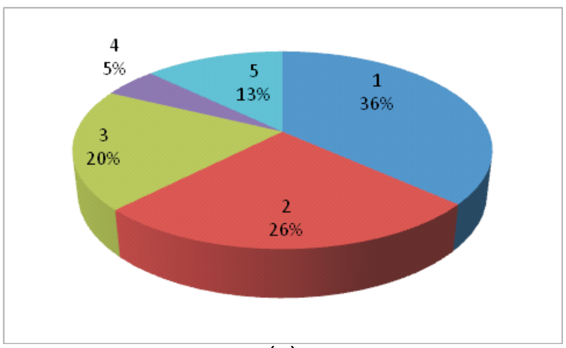
4. We analyzed the fingerprint images specifically those which are causing errors. As shown in Figures 4-6, we found that there are some specific causes that are more relevant in Indian sub-continent compared to western and European countries. Lawsonia Inermis (commonly known



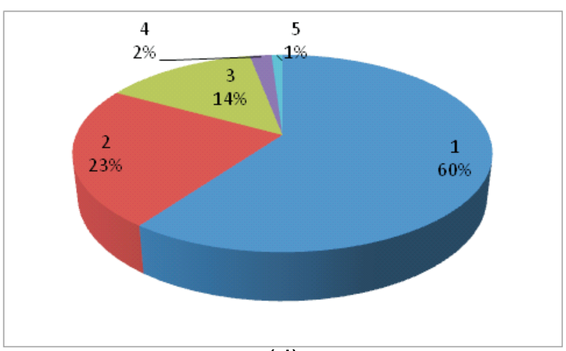

(a)




(b)




(c)




(d)

**Figure 1. Illustrating the NFIQ score distribution in (a) DB1-Urban, (b) DB1-Rural, (c) DB2, and (d) DB3. In the pie-charts, each block with number 1-5 represents NFIQ scores. Note that NFIQ scores range from 1 (excellent) to 5 (poor).**

**Figure 2. An example where a segmentation tool failed to correctly segment the images.**

as heena or mehandi) can cause significant differences in quality of fingerprint images. Widely used by women in Indian sub-continent during festivals, heena is applied on hand/fingers and when applied, fingerprint sensors may not be able to properly capture fingerprint information. Further, patterns obtained from individuals belonging to labor class, often show a quality score of 4 or 5. Visual analysis suggests that such images have worn and damaged patterns and therefore recognizing such patterns would be difficult (Figures 5 and 7).

5. On analyzing the quality distribution of each finger in every age group, it is difficult to generalize that little fingers are useful or not. Similarly, we cannot generalize that, a particular age group or gender conforms to lower or higher quality scores and hence better/worse performance.

6. We observe that some simple operational processes such as

    (a) using wet towels to remove dirt and dry fingers,

    (b) using minimum quality threshold to ensure that extra efforts are made to capture good prints from hard to obtain fingers and

    (c) keeping scanning devices in operational order

    can help in achieving better quality images and therefore improve the accuracy

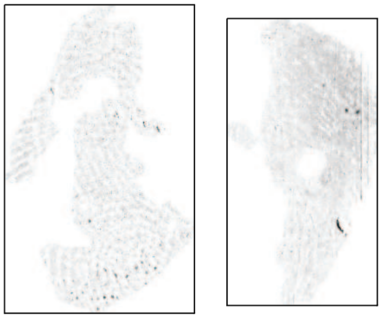


**Figure 3. An example where NFIQ incorrectly assigned a quality score of 1 and 2 to a poor quality image.**

**Figure 4. Difference in image quality can cause reduced performance: 35 years old housewife in rural settings (second image with heena/mehandi applied on the finger).**

**Figure 5. An example illustrating a case when it is difficult to capture good quality fingerprint images.**

**Figure 6. There are some features such as scars and warts that are difficult to analyze in current algorithms/systems.**

In another experiment, we use DB1 (both urban and rural components) to compute Rank-10 identification accuracy (DB2 and DB3 do not have multiple samples per finger and therefore genuine-impostor match score analysis is not possible). Two commercial softwares were used for enrollment and identification. We selected one sample per finger as gallery and the remaining as probe. Further, 10 times cross validation is performed using the gallery-probe partitioning. Identification accuracies in Table 1 summarize the results of this experiment. Experiments with DB1-urban fingerprints show that both the systems yield higher rank-10 identification accuracy. We observe that with high quality images, systems provide accurate results whereas the performance suffers when the fingerprint quality is poor or with limited number of minutiae (i.e. reduced region of interest). On the other hand, experiments with rural fingerprints show remarkable results. Since the database contains several non-ideal poor quality fingerprint images, accuracy of the systems reduce significantly. On this database, best accuracy is less than 65% which is certainly not acceptable in large real world applications. Decrease in performance is mainly because of the poor quality images which cause either missing or spurious fingerprint features. We performed match score-quality score correlation analysis and observed that quality scores are highly correlated with the match scores (correlation of 0.82). The correlation analysis suggests that the image quality can be a good indicator of performance especially when match scores are not feasible to compute (as in case of DB2 and DB3). We therefore applied this analysis on DB2 and DB3 databases that have only single impression per finger. Using existing western results (i.e. NIST IR 7151: correlation of match scores with NFIQ quality scores), we can closely predict the expected fingerprint recognition performance in Indian environment. In the experiments, we observed that, at 95% confidence, DB2 is expected to show lower accuracy compared to western data whereas DB3 is expected to achieve similar accu-

**Figure 7. An example where ridge patterns are worn and damaged.**

**Table 1. Rank-10 identification accuracy of different fingerprint recognition systems.**

| System | Urban | Rural |
|---|---|---|
| System - 1 | 90 | 64 |
| System - 2 | 91 | 64 |

racy. However, to accurately predict the accuracy, it is required to collect multi sample database from rural Indian population.

## 4. Conclusion

This paper presents an image quality based feasibility study for fingerprint biometrics in Indian environment. This experimental analysis is important since there is no prior study that provides any insight about using fingerprints in large scale applications such as UID. The study suggests that fingerprints pertaining to Indian population requires additional research to address the region specific problems such as worn and damaged patterns.

## 5 Acknowledgement

The authors would like to acknowledge UIDAI and biometric standards committee members for their help and support.

## References

[1] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar. *Handbook of fingerprint recognition*. Springer Verlag, 2003.

[2] NFIQ. http://fingerprint.nist.gov/nbis/index.html, Last accessed on Jan 25, 2010.

[3] U. I. A. of India. http://www.uidai.gov.in/, Last accessed on Jan 25, 2010.