Does Habituation Affect Fingerprint Quality?

Mary Theofanos
National Institute of Standards and Technology
100 Bureau Drive, MS 8940
Gaithersburg, MD 20899
mary.theofanos@nist.gov

Ross Micheals
National Institute of Standards and Technology
100 Bureau Drive, MS 8940
Gaithersburg, MD 20899
ross.micheals@nist.gov

Jean Scholtz
National Institute of Standards and Technology
100 Bureau Drive, MS 8940
Gaithersburg, MS 20899
jean.scholtz@nist.gov

Emile Morse
National Institute of Standards and Technology
100 Bureau Drive, MS 8940
Gaithersburg, MD 20899
emile.morse@nist.gov

Peter May*
George Washington University
2121 Eye St
Washington, DC 20052
psmay@gwu.edu

CHI 2006, April 22–27, 2006, Montreal, Quebec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract
Interest in the environmental factors that affect biometric image quality is increasing as biometric technologies are currently being implemented in various business applications. This study aims to determine, through repeated trials, the effects of various external factors on the image quality and usability of prints collected by an electronic reader. These factors include age and gender but also the absence or presence of immediate feedback. A key factor in biometric systems that will be used daily or routinely is habituation. The user's behavior could potentially change as a result of acclimatization; one's input might increase in quality as one learns how to use the system better, or decrease in quality since comfort with the system could translate into carelessness.

Keywords
Habituation, Fingerprint image quality, feedback, biometrics

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g., HCI)]: User Interfaces. General Terms: Human Factors

Introduction
Interest in biometric access control—the use of one or more of a person's inherent biological properties (fingerprints, iris, hands and faces, *et cetera*)—is rapidly growing within the government as well as the private

User Interface

Figure 1. Screenshot of the user interface with the quality 'gauge' expanded.

sector. Biometric access has attractive advantages over other modes of access control. Unlike token-based control (e.g., physical key access), there is no physical artifact to misplace or to have stolen, and unlike knowledge-based control (e.g., password access), there is no secret phrase to forget or to pass on to an unauthorized party. Biometrics are best used as a component in a broad security plan. Therefore, many factors must be considered when selecting both the biometric and the type of application in which the biometric will be used.

A taxonomy of biometric applications introduced by Wayman [1] identifies seven categories by which applications can be categorized: ① *cooperative* v. *non-cooperative* users, ② *overt* v. *covert* capture, ③ *habituated* v. *non-habituated* users ④ *attended* v. *non-attended* (if the system requires an assistant) ⑤ *standard* v. *non-standard environment* (e.g., indoors at an airport or at a border crossing), ⑥ *public* v. *private* (if users include the general public or not), and ⑦ *open* v. *closed* (if the system interoperates with external systems).

Each of these categories has specific challenges associated with it and requires further investigation since adoption of biometric applications depends on performance of the system. In this study we assumed a cooperative, overt environment where users would be habituated. We also used a standard indoor environment. No assumptions were made about the public versus private or open versus closed aspects of the system.

We selected electronic fingerprint collection as a domain for investigating habituation, since the use of fingerprints as identification is already well-established. Our study objectives were to determine if habituated users gave better quality prints than non-habituated users and if feedback affected the quality. Studies carried out by National Institute of Standards and Technology (NIST) have found that a significant amount of variation in system performance may be attributed to fingerprint image quality [2]. Investigation of fingerprint image quality [3, 4] found that factors such as age, gender, and skin conditions affect system performance and anecdotally image quality of a fingerprint system.

The Government and other organizations are planning to install fingerprint readers to allow employee access to various parts of buildings. This may require an employee to submit biometric input a number of times per day. Does habituation affect the user's performance and the acquisition of quality prints? The user's behavior could potentially change as a result of acclimatization; one's input might increase in quality as one learns how to use the system better, or decrease in quality since comfort with the system could translate into carelessness. And if we find differences in behavior over time, when does a user become habituated?

Additionally we are interested in how feedback affects habituation and image quality. We addressed this in the second phase of our study. There are many types of feedback to consider. For most biometric systems, the granularity of feedback presented to the user is typically quite coarse. For example, a fingerprint-based access control system might only include feedback indicating when a subject should start and stop presenting their biometric, or indicate if a claim of identity was accepted or rejected (e.g., door opens or remains closed). But there are applications such as US-VISIT in the airport that may or may not accept the print. In this case how does the traveler know if the print was rejected due to the quality of the print or some other issue? What type of feedback is appropriate if the print was rejected due to the poor quality? What information helps the user improve the fingerprint image quality?

Fingerprint Images

Figure 2. Fingerprint images of high (a) and low (b-c) quality. Because NFIQ is rank based, lower values correspond to higher quality.

(a) NFIQ=1

(b) NFIQ=5

(c) NFIQ=5

Experiment Design & Implementation

The study was divided into two phases. Prior to each phase each participant completed a questionnaire that included demographic data such as age, gender, and education as well as questions on security and privacy issues. In addition all users were asked to select a login pass phrase that they used to access the finger print reader. This allowed us to keep the data anonymous but ensure that we could accurately attribute the fingerprints and the demographics to the proper user.

During the first phase we collected a left and right index finger image from participants. Participants were encouraged to leave one set of prints before lunch and one after lunch as they entered and left the cafeteria over a three week period. We encouraged participants to leave at least 20 images during that period of time. During this phase participants were not allowed to view the images and were given no feedback as to the quality of the print. Thus the operator indicated to the participant which finger to place on the scanner and when the participant should remove her finger. In addition, the operator viewing the image on the screen manually captured the image. The operator did not make a decision about quality – but merely accepted a fingerprint once the image stabilized on the screen.

In the second phase we also collected left and right index finger images. Again, participants were encouraged to leave a set of prints before lunch and after lunch as they entered and left the cafeteria over a 3 week period. As in the first phase they were encouraged to participate at least 20 times during that period of time. But this time the participants were provided "real-time" feedback as to the quality of the fingerprint image and they determined which image to save. Through the user interface participants indicated which index finger to scan. Once they were comfortable with the positioning of their finger on the scanner, they captured the image. A fingerprint quality score (NFIQ, discussed in the next section) was returned and displayed on a dynamic scale bar. Figure 1 shows a screenshot along with a detailed view of the expanded quality gauge. A 'full' gauge was considered 'best.' Participants were encouraged to collect as many samples as they desired in order to capture an image with an NFIQ score of 3 (borderline) or higher. Once satisfied with their score participants submitted the images to the database. We saved every attempt in the database and not just the final image. We wanted the feedback to be as simple as possible so that participants did not have to interpret the feedback—such as showing them the print.

Fingerprint background

Before discussing the results of our study, we must (a) further define image quality and (b) distinguish between manual versus automatic capture. Their interaction had a significant impact on our experimental design.

Image quality. As discussed previously, there is a variety of research on the nature of fingerprint image quality. In this study, image quality is defined and measured in accordance with the NIST Fingerprint Image Quality (NFIQ). NFIQ defines quality "as a predictor of matcher performance before a matcher is applied." Taking a single fingerprint image as input, NFIQ outputs an integer value from one to five (inclusive) where images of quality one are the most likely to yield correct results. NFIQ software is freely available from NIST (but export controlled).

Figure 2 shows a variety of fingerprint images along with their accompanying NFIQ scores. Notice that images that we might observe to be visually pleasing do not always map to fingerprints that are predicted o give accurate results. Specifically, image (c) maintains many clear edges and a superior contrast as compared to (b), yet they share the same NFIQ value.

Since there can be an apparent disconnect between what *appears* to be a high-quality image and what is *measured* by NFIQ as a quality image, we decided not to use the fingerprint image as feedback. Instead we present the NFIQ quality score on a gauge.

Manual v. Automatic Capture. In a manual capture scenario, a fingerprint sensor behaves much like a video camera; outputting a stream of images at some frame rate. At some explicit trigger, a particular frame (either the most recent or the very next) is captured and used as the input to some other system component. With automatic capture (or *auto capture*), an explicit or implicit trigger does not immediately capture an image, but instead slaves the sensor into a polling mode. Then, it is the responsibility of the sensor to return a fingerprint image within some timeout period. Naturally, there is a wide variety in the amount and sophistication of processing that a sensor could perform in order to make this determination. For example, a naïve sensor might simply wait until a certain number of pixels are active. A more robust sensor may only return an image in which a sufficient number of minutiae (fingerprint features) or a particular quality level is detected.

Because there is no generally accepted best-practice for the implementation of automatic capture, we opted for manual capture with immediate computation and presentation of the quality score.

Phase One Results. Twenty-one men and 17 women volunteered to participate in the study. The 12 men and 17 women who eventually participated were middle-aged on average but the group had members in their teens as well as those in their 70s. As shown in Figure 3, we found that younger subjects submitted higher quality prints than older subjects. The age-related phenomenon was true for both men and women. This confirms the results of other small scale studies [4], but formal analysis of the image quality across a large population has yet to be performed. In addition, the women in our sample submitted fingerprints that were, on average 20% poorer in quality than the men in the group.

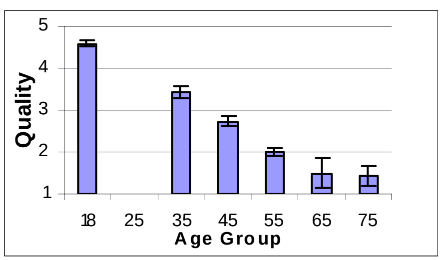

Figure 3. Quality of fingerprints by age group

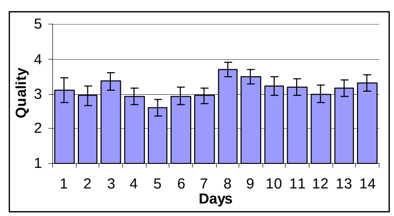

Figure 4. Quality of finger prints over time

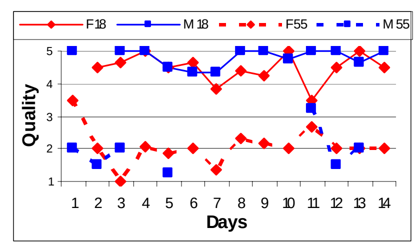

Figure 5. Quality of finger prints over time for 18-25 and 55-65 age groups

As evidenced by Figure 4, over the course of the 14 days of print collection there did not appear to be any habituation effect. This graph is an aggregate of the quality scores across all of the test participants. The variability is seen more clearly in Figure 5 which focuses on four groups of participants. This figure depicts the day to day values of the male and female participants in the groups with the largest participation – those in the 18 to 25 and 55 to 65 age groups. While both groups demonstrated a great deal of variability daily they never overlapped. Once again the younger subjects submitted consistently better quality prints than the older subjects. While print quality differed from day-to-day for individuals, there was no overall trend toward significantly higher or lower quality prints over time.

Phase Two Results. In the second phase, we recruited 24 men and 16 women. However, only 17 men and 11 women actually participated. Again, the men and women who eventually participated were middle-aged on average but the group had members in their teens and in their 60's. Again, as shown in Figure 6, we found that younger subjects submitted higher quality prints than older subjects.

Since phase two provided "real-time" feedback as to the quality of the fingerprint image and participants determined the number of attempts to produce a good quality image; we also computed the number of attempts by age group. This is shown in Figure 7. Not only was the quality of prints overall worse for the older age groups, but they tried more times to produce a good print than did the younger participants. Figure 8 shows that the youngest participants submitted high quality fingerprints throughout the study. There was no habituation effect for this group; the only direction it could have gone was down. This is consistent with Phase I results which showed that people in the 18–25 age range give consistently good prints. The dotted lines in the figure show the quality of prints from men and women in the 55–65 year age range. In Phase I, we found that these people produced borderline prints. The figure shows that older subjects produced prints that were of higher quality over the course of the study, and surprisingly, women out-performed men. By the middle of the study period, the older women submitted prints that were of the best quality. In addition, it took fewer attempts to capture a good quality print towards the end of the study.

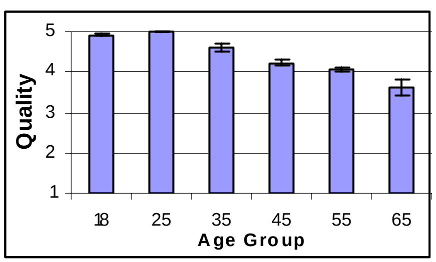

Figure 6. Quality of fingerprints by age group

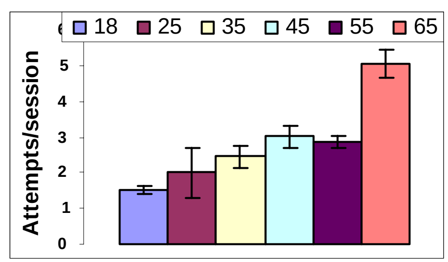

Figure 7. Attempts by age group

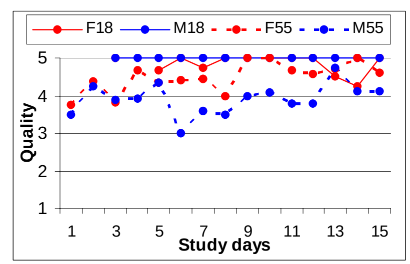

Figure 8. Quality of finger prints over time for 18-25 and 55-65 age groups.

We observed that the older subjects were very determined to obtain a good quality image. Since the feedback provided did not relay detailed information on the quality beyond the score, the user didn't know how to reposition her finger, or whether to press harder or softer, or if her finger was too dry. This was frustrating to many users. But many saw this as a challenge and became competitive. This led to the development of strategies and techniques that would result in capturing a good quality image.

Conclusions

We observed three main results. First, we confirmed the general consensus of the biometrics community studying the NFIQ quality measure that age and gender do affect quality. People in the 18–25 age range give consistently good prints, while older individuals have more borderline print quality. In general, men give higher quality prints than women.

Second, habituation with no feedback at all was not shown to affect the quality of prints. Print quality differed from day-to-day for every individual. There was no overall trend toward significantly higher or lower quality prints over time.

Finally, feedback did translate into improvement of quality. By the end of the study period, in general subjects produced higher quality prints with fewer attempts. However participants did not know what to do to correct poor quality prints. Thus finer grained feedback may be required. It is difficult to provide good feedback since so many variables affect quality including the position of the finger on the scanner, pressure, and moisture. How do you indicate which variable needs to be addressed without confusing the user? Even providing the image itself presents problems for participants, since only experienced experts might look at a print and indicate the quality and interpret the reason for poor quality.

The above factors afford many possibilities for future investigation of feedback for biometrics input. We must also point out that this work was done using the assumptions of a cooperative, overt, standard operating environment. Changing any one of these assumptions would require a new study.

*Sponsored by the Summer Undergraduate Research Fellowship (SURF) program, a partnership between the National Science Foundation (NSF) and NIST.

References

[1] Wayman, J., Jain, A., Maltoni, D., and Maio, D. *Biometric Systems Technology, Design and Performance Evaluation.* Springer-Verlag, London, England, 2005.

[2] Tabassi, E., Wilson, C., and Watson, C. Fingerprint Image Quality, NISTIR 7151, August 2004, http://www.itl.nist.gov/iad/894.03/fing/fing.html.

[3] Maltoni, D., Maio, D., Jain, A., and Prabhakar, S. *Handbook of Fingerprint Recognition*. Springer-Verlag, London, England, 2003.

[4] Elliott, S.J., Kukula, E.P., and Sickler, N.C. The Challenges of the Environment and the Human/Biometric Device Interaction on Biometric System Performance, International Workshop on Biometric Technologies-Special forum on Modeling and Simulation in Biometric Technology, Calgary, Alberta, Canada, 2004.