Analysis of Slap Segmentation and HBSI errors across different force levels

Adam Wamsley
Biometric Standards, Performance and Assurance Laboratory
Purdue University, Department of Technology, Leadership and Innovation
West Lafayette, IN

Stephen Elliott, Ph.D.
Biometric Standards, Performance and Assurance Laboratory
Purdue University, Department of Technology, Leadership and Innovation
West Lafayette, IN

Carl Dunkelberger
Biometric Standards, Performance and Assurance Laboratory
Purdue University, Department of Technology, Leadership and Innovation
West Lafayette, IN

Mitch Mershon
Biometric Standards, Performance and Assurance Laboratory
Purdue University, Department of Technology, Leadership and Innovation
West Lafayette, IN

***Abstract*— In a recent study, fingerprint data was collected across six different force levels. A total of 8,877 slap samples were ground truthed, and subsequently processed using a commercially available fingerprint segmentation tool. This paper delves deeper into understanding segmentation errors across the force levels. Out of the 8,877 slaps, 370 slaps failed to segment. In order to understand why there were errors, a detailed analysis was undertaken by ground-truthing the slap and segmented datasets. In addition to the ground-truthing, video analysis of the interactions enabled specific failures to be replayed and identified as performance or ergonomic, interaction, and usability (Human Biometric Sensor Interaction) errors. This paper will identify determining factors that would cause the slap segmentation tool error, and provide guidance to those undertaking data collection activities where they would need to the use a segmentation tool.**

Keywords-Fingerprint recognition; image quality; law enforcement; biometrics

I. Introduction

According to the UIDIA "it is very important to pay attention to the ergonomics of the set up to get the best possible biometrics. An important factor in obtaining better image quality is operational process" (pg 69) [1], and in an operational setting, ensuring that a correct image has been taken from the subject is key in optimizing the performance of a biometric system. This paper examines whether the interaction between the human and the sensor contributes to segmentation error. In order to answer this question, a number of slap images were segmented and classified. Those returning an error were further processed by watching videos of the interaction captured during collection. This methodology is part of the Human Biometric Sensor Interaction model [2], and combined with video analysis, provides detailed interaction data that can be used post-hoc to illustrate issues in subject interaction. With the ever growing need for optimal throughput at security locations manually segmenting fingerprints will act as a bottleneck and may even deter some from using the system. It is hoped that this research will also provide guidance for those training data collection agents to optimize their process in an operational or laboratory setting, as well as understanding the types of interaction that subjects could potentially do. By doing this experiment those errors caused by segmentation and human biometric sensor interaction can be identified and resolutions made to correct.

There are many different factors can cause a disruption in image quality such as age, occupations with abrasion and chemical exposures, amputations, scars, skin diseases, improper finger placement from orientation to the tip of the finger, scanner limitations, environment conditions such as temperature and humidity, and unknown substances on scanner surface [3-6]. In addition, poor interaction combined with the above factors can accesserbate the degradation in performance. For this experiment the term biometrics is being looked at as a process and a characteristic. As defined by NSTC "Biometrics Glossary" biometrics as a characteristic is "a measurable biological (anatomical and physiological) and behavioral characteristic that can be used for automated recognition" and as a process "biometrics refers to automated methods of recognizing an individual based on measurable biological (anatomical and physiological) and behavioral characteristics" [7]. Slap segmentation is done by dissecting a slap image into four individual fingerprint images [8]. This is normally done after a successful slap has been presented to a scanner. A successful slap is defined as four fingerprints of high image quality or three fingerprints of high image quality and correct fingerprint position identification. Segmentation of images can be done by either real-time segmentation, when the images are immediately segmented after capture or by batch segmentation, when the images are segmented from a previously collected dataset. For this experiment batch segmentation was performed on an existing dataset.

The Human Biometric Sensor Interaction (HBSI) model has been developed over the past six years, with the intent of understanding successful and erroneous interactions with the biometric sensor related to behaviors, movements, and other human actions. These erroneous interactions would originally not have been documented or would have been falsely attributed as other errors. The model has gone through several iterations as the state-of-the-art has progressed from single instance to “smart” continuous capture sensors. The research underlying the model can be found in [9-12] and was developed using a single instance or event based biometric system. An example of an event based biometric system is a swipe fingerprint system where the subject swipes their finger across the sensor and the device responds either by accepting or rejecting the single interaction (event). There are five different errors, stemming from the decision of an incorrect or correct presentation. Exploring the incorrect presentation component of the model, three different errors are presented: defective interaction (DI), concealed interaction (CI), and false interaction (FI). There are two main challenges with conducting HBSI evaluations. First, researchers must understand the concept of operations for a particular device and identify all possible errors that could occur between the human, sensor, and system. Second, consistent and correct classifications of the potential errors according to the parameters outlined must occur.

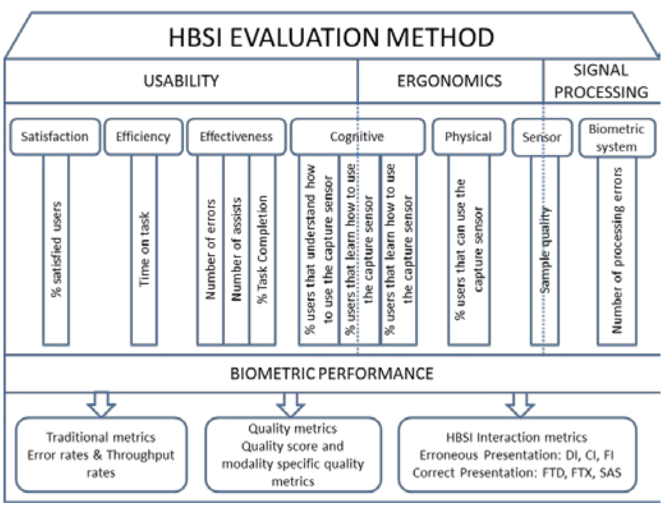

Figure 1: HBSI Evaluation Methodology

The HBSI evaluation (Figure 1) illustrates the connection between usability, ergonomics and signal processing. As part of this study, the effectiveness of the ten-print fingerprint scanner is measured – using the following metrics: number of errors (in this case the number of errors from the segmentation tool), the number of assists (calculated from the prompts from the test administrator) and the % of task completion (in this case the ability of the subjects to successfully cause the device to “beep” that it has acquired the image.

II. METHODOLOGY

The underlying research question is to understand whether interaction errors contributed to segmentation errors, using publically available commercial segmentation software, which provided an output of 15 metrics. To code the specific interaction errors, there were three video recording of the interaction. The first view facing down onto the hand / sensor, the second was a side angle of the interaction, and the final feed was the screen that was seen by the test administrator. All three screens were synced and simultaneously recorded for subsequent analysis. The selection of the camera angles and locations was driven by the research question which was to evaluate the effect of force on fingerprint image quality, however, each of these screens collected information about the interaction; thereby enabling a secondary and systematic review of the data based around segmentation error.

The software used to categorize the various behavioral observations is the Observer XT version 10 from Noldus Inc. The software is an automated system for coding behaviors and enables researchers to answer research questions that relate to behaviors. There are three classifications in the software; subjects, behavior and modifiers. The observational methodology is as follows: firstly all of the videos were observed and a list of potential errors was reviewed. These were then classified as behaviors. Due to the short observational period for each interaction, the observations were coded as point events. Such events are behaviors that are observed, rather than have a duration associated with them. The behavioral measures are coded as a frequency. The following behaviors were measured: during placement of the slap, the following two different entry methods occurred - firstly the subject can place their hand where all fingers are flat, or where the fingers come in one at a time, which is analogous to a "piano playing" movement. The next set of movements occur once the subject has placed their hand on the platen; this could be correction of the placement, and this can be prompted by either the operator (via a spoken command), or by the subject not placing their hand properly on the sensor. This may occur if their index finger or little finger does not place the first time on the sensor. The second movement measured is a stable hand - where the subject does not move at all, but does not reach the appropriate force level. The third movement is where the subject's handshakes on the platen. We will align this shaking with self-disclosed survey question relating to musculoskeletal diseases. The last component of the behavioral measurement relates to the release of the hand from the sensor. This can be done in the following ways - either after acquisition of the image (coded by the beep of the sensor), a premature lift, or a timeout. The timeout occurs after 20 seconds. Each of the segmentation errors were analyzed by reviewing the video interaction, the distribution of these errors is illustrated in the results section.

In concert with the video interaction, each slap image was ground truth, both visually as well as algorithmically. This ensured that the data was synchronized with all the associated metadata.

III. RESULTS

The dataset consisted of 8,877 slap images from 247 individuals, which were collected using the Crossmatch LScan Guardian livescan device, at five different force levels 5N, 7N, 9N, 11N, and 13N. In addition to these five force levels, images were also collected using auto capture setting on the Crossmatch device.

A. General Results

During the collection process all force levels except for auto capture were randomized for each subject. In the larger study, 259 subjects were collected, and 247 subjects were analyzed. The differences can be attributed to poor synchronization in the video feeds, or data collection errors identified as part of the ground truthing process. The total number of interactions was 72, resulting in 43,723 fingerprint images (post segmentation). For the purposes of this paper, we provide the total number of errors in terms of the number of interactions. Out of the 247 subjects, 136 exhibited some form of segmentation error, resulting in 370 interactions that prompted an error. Out of those interactions, 164 were attributed to males, 206 were attributed to females, and 193 errors were associated with the left hand, 177 to the right hand. The distribution of hands that caused errors (left / right) were not correlated with the subject's dominant hand.

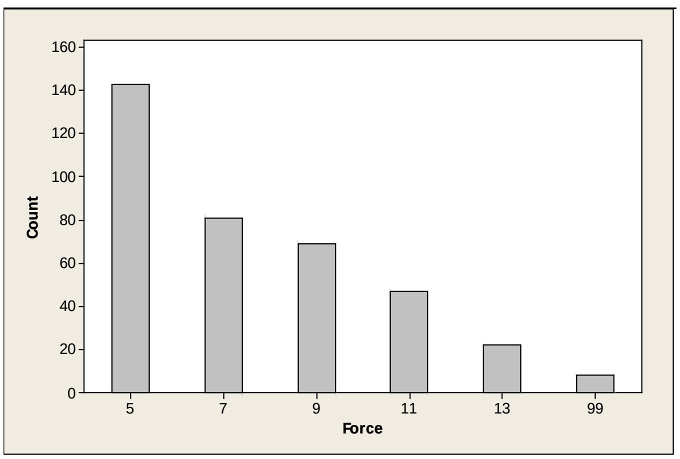

Figure 2: Segmentation Errors by Force

Figure 1 shows the distribution of segmentation errors by force, with "99" denoting auto capture. Segmentation errors also increased with the number of trials (from 107 for interaction #1 to 134 interaction #3).

B. Segmentation errors grouped by force levels.

The next stage of the analysis was to examine the individual aspects of the segmented errors, and understand whether there were any statistical differences across force levels. There was no statistically significance difference in slap orientation versus force ($p=0.588$). There was no significant difference on digits found by the segmentation tool and hand, however, there was a statistically significant difference in the number of digits found and force levels. This correlates with the findings that there were significant differences in image quality across the different force levels, with the lower force levels (5N) having significant difference than the remaining force levels. With respect to finger detection, Figures 3 and 4 indicate the number of digits found by force. Figure 4 shows that auto capture (0) was the best in finding fingers, followed by 13, 11, 9, 7 and 5N respectively.

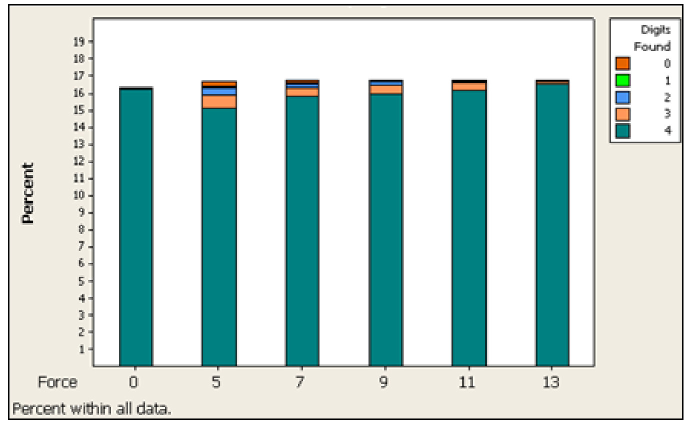

Figure 3: Number of Digits Found by Force

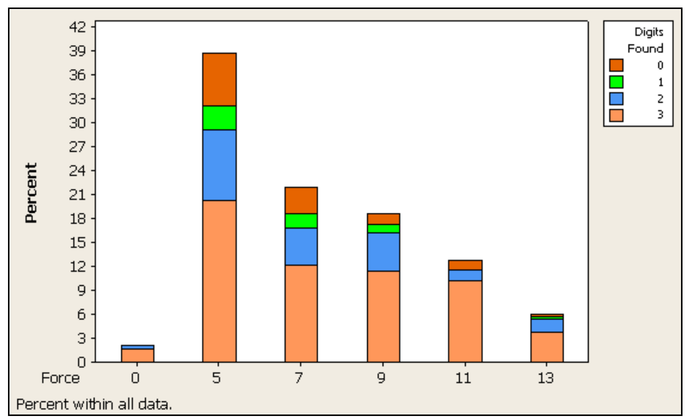

Figure 4: Number of 3 or Less Digits Found by Force

C. Interaction results

In order to establish whether there were any interactions that could cause these segmentation issues, we divided the capture process into three parts – the entry, the capture, and the exit. In each interaction, the method of entry was coded.

1) Entry methods

Ninety percent of the entry errors were defined as flat – where each of the fingers was lined up in a plane. The piano entry method, where each finger was presented to the platen in sequence (akin to playing a piano) accounted for only 8%. The remainder (2%) could not be coded in either the flat or piano entry category. There is no statistically significant difference in the entry method versus image quality for the hand, or separately, for the fingers. When modifiers are taken into consideration, there is a statistically significant difference in image quality across the modifiers (finger tips, fingers bent up, no obvious placement, and pivot), with fingertips, no obvious placement and pivot being statistically significantly different than fingers bent up. The image quality mean for fingers bent up was 9.5 (on a scale of 1-100). Table 1 shows the distribution of modifiers on the placement of the hand. Although the majority of the flat entries had no obvious placement, 237 of the segmentation errors were attributed to this. Twenty-five percent of the errors can be attributed to fingertip entry.

Table 1: Modifiers on Placement of Hand

Modifiers on Placement of Hand	Number of errors	%
Flat / Fingers Bent	18	4.86%
Flat / Pivot	33	8.91%
Flat / Fingertips	82	22.16%
Flat / No obvious	237	64.05%

2) Hands on the sensor

The second area of investigation was how the hands acted when placed on the sensor. Here the choices were no movement or movement with the appropriate modifiers as sufficient force or insufficient force. The majority of interactions had no movement, and this was not significant across force levels. Out of the 370 interaction errors, only 24 interactions included task correction, with 16 of those not being prompted. It should be noted that the subjects could not see the image of their fingerprints on any screen, and therefore they made the judgment to move their hands. Follow on research will examine the non-error group to establish the frequency in which that group task corrected.

3) Removal of hand from the sensor

The majority of hands (over 99%) removed their hand at the appropriate time. The remainder prematurely lifted their hand. Due to the small sample size, no determination of the general cause of this can be made.

D. HBSI Evaluation Metrics

The HBSI evaluation metrics that measure effectiveness are the number of errors from the segmentation tool, the number of assists, and the percentage of task completion. In this study, the number of assists and the percentage of task completion refer only to the error dataset, not the entire 8,877 slaps. The number of errors from the segmentation tool was 4.17%. The number of assists, defined as intervention from the test administrator on only the errors was 2.16%. The % of task completion was 99%.

IV. Conclusions and Recommendations

It is interesting to note that the majority of errors in the segmentation software program came from interactions that exhibited flat entry, no obvious entry characteristics, no movement on the sensor, and sufficient force – interactions that were not out of the ordinary expectation in the use of the sensor. Furthermore, it is also interesting to note that the method of entry is not impacted by force level. A recommendation that should be made when training individuals is that they should not bend their fingers up during data collection. This results in very poor image quality. Conversely, it is also interesting to note that there finger tips do not contribute a large number of errors.

V. References

[1] *General Guidelines to Enrolment Agencies - Guidelines for setting up Enrolling Centres*, 2010.

[2] E.P. Kukula, M.J. Sutton, and S.J. Elliott, “The Human–Biometric-Sensor Interaction Evaluation Method: Biometric Performance and Usability Measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, Apr. 2010, pp. 784-791.

[3] C.R. Blomeke, S.K. Modi, and S.J. Elliott, “Investigating the relationship between fingerprint image quality and skin characteristics,” *2008 42nd Annual IEEE International Carnahan Conference on Security Technology*, Prague, Czech Republic: IEEE, 2008, pp. 158-161.

[4] Z. Li, Z. Han, and B. Fu, “A Novel Method for the Fingerprint Image Quality Evaluation,” *2009 International Conference on Computational Intelligence and Software Engineering*, IEEE, 2009, pp. 1-4.

[5] E. Lim, K.-ann Toh, P.N. Suganthan, J. Xudong, and Y. Wei-Yun, “Fingerprint image quality analysis,” *2004 International Conference on Image Processing, 2004. ICIP ’04.*, Ieee, 2004, pp. 1241-1244.

[6] C. Jin, H. Kim, X. Cui, E. Park, J. Kim, J. Hwang, and S. Elliott, “Comparative Assessment of Fingerprint Sample Quality Measures Based on Minutiae-Based Matching Performance,” *2009 Second International Symposium on Electronic Commerce and Security*, Nanchang, China: IEEE, 2009, p. 309–313.

[7] “Biometrics Overview,” *Biometrics*, 2006, pp. 1-10.

[8] A. Hicklin and R. Khanna, “The role of data quality in biometric systems,” *White Paper. Mitretek Systems (February 2006)*, 2006.

[9] K. Eric P. and E. Stephen J., “Implementing Ergonomic Principles in a Biometric System: A Look at the Human Biometric Sensor Interaction (HBSI),” *Proceedings 40th Annual 2006 International Carnahan Conference on Security Technology*, Lexington, KY: IEEE, 2006, pp. 86-91.

[10] E.P. Kukula, M.J. Sutton, and S.J. Elliott, “The Human–Biometric-Sensor Interaction Evaluation Method: Biometric Performance and Usability Measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, Apr. 2010, pp. 784-791.

[11] E.P. Kukula and S.J. Elliott, “Ergonomic Design for Biometric Systems,” *Encyclopedia of Biometrics*, 2009, pp. 274-280.

[12] E. Kukula, S. Elliott, and V. Duffy, “The effects of human interaction on biometric system performance,” *First International Conference on Digital Human Modeling (ICDHM 2007), Held as Part of HCI International*, Beijing, China: Springer, 2007, pp. 904-914.