Effects of JPEG 2000 Lossy Image Compression on 1000 ppi Latent Fingerprint Casework

Shahram Orandi
John M. Libert
John D. Grantham
Frederick R. Byers
Lindsay M. Petersen
Michael D. Garris

Effects of JPEG 2000 Lossy Image Compression on 1000 ppi Latent Fingerprint Casework

Shahram Orandi
John M. Libert
Michael D. Garris
Frederick R. Byers
Information Access Division - Image Group
Information Technology Laboratories

John D. Grantham
Systems Plus, Inc.
Rockville, MD

Lindsay M. Petersen
MITRE Corporation
McLean, VA

October 2013

The criminal justice communities throughout the world exchange fingerprint imagery data primarily in 8-bit gray-scale and at 500 pixels per inch[1] (ppi) or 19.7 pixels per millimeter (ppmm), and these are exchanged among agencies in a "compressed" format. Compression is a method whereby information is encoded in a way that requires fewer bits than the original representation requiring less storage provisions or allows for transmission through a network faster. Compression can be employed using one of two approaches, either lossless or lossy. In lossless compression, the resulting compressed data represents the original data with no loss in fidelity. Lossless encoding typically can reduce the relatively complex data into half its original size. Lossy compression algorithms on the other hand employ data encoding methods which discard (lose) some of the data in the encoding process in order to achieve more aggressive reduction in the size of the data being compressed over what is possible with lossless methods (on the order of 5 to 10 times better reduction in size versus lossless methods).

Exemplar fingerprints are fingerprints collected using controlled measures (for example during the booking process) while latent fingerprints are captured in an unconstrained/uncontrolled capture means such as contact surfaces at a crime scene. Since exemplar fingerprints are produced through controlled capture measures, they typically are of better quality than latent fingerprints.   Because of this quality difference, the prescribed method for the compression of exemplar fingerprints has been with lossy compression methods while latent fingerprints have been typically compressed using lossless methods.

The Wavelet Scalar Quantization (WSQ) fingerprint image compression algorithm is currently the standard algorithm for the compression of 500 ppi exemplar fingerprint imagery. WSQ is a "lossy" compression algorithm. Decompressing the resulting compressed data yields content that, while different from the original, is similar enough to the original that it remains useful for the intended purpose. The WSQ algorithm allows for users of the algorithm to specify how much compression is to be applied to the fingerprint image at the cost of increasingly greater loss in fingerprint image fidelity as the effective compression ratio is increased (see Figure 4 on page 18 for an example of image degradation from lossy compression). The WSQ Gray-Scale Fingerprint Image Compression Specification [WSQ] provides guidance based on an International Association for Identification (IAI) study [FITZPATRICK] to determine the acceptable amount of fidelity loss due to compression in order for a WSQ encoder and decoder to meet FBI certifications. These certifications are designed to ensure adherence to the WSQ specification and thereby to ensure fidelity and admissibility in courts of law for images that have been processed by such encoders and decoders.

In 2011, NISTIR-7778 [NIST2] built upon the IAI study with expanded scope for rolled and flat (exemplar) fingerprint imagery at 1000 ppi compressed using the JPEG 2000 [JPEG2k] compression algorithm. NISTIR-7778 showed that a target compression rate of 10:1 may be optimal for exemplar images (fingerprints obtained using constrained/controlled fingerprint capture means). NISTIR-7778 did not, however, examine the impact of lossy compression on exemplars when used for comparing with latent fingerprint imagery (fingerprints obtained from unconstrained/uncontrolled capture means such as contact surfaces at a crime scene), which is the focus of this study.

This study (NISTIR-7780) examines the effects of lossy compression using the JPEG 2000 algorithm in terms of image degradation as observed by examiners as well as the effects of increasing compression on the ability of examiners to make their identification decisions.

For two key scenarios using 1000 ppi images that best represent typical operational latent casework where a lossless (never processed by lossy compression) latent fingerprint is compared to either a rolled or flat lossy compressed exemplar fingerprint to establish identity, the maximum exemplar compression ratio that yields images that are indistinguishable in terms of subjective evaluation by latent examiners versus a lossless 1000 ppi exemplar is 12:1 (see Table 13).

---

[1]Resolution values for fingerprint imagery are specified in pixels per inch (ppi) throughout this document.   This is based on widely used specification guidelines for such imagery and is accepted as common nomenclature within the industry. SI units for these will be presented only once.

While this study confirms that even at the highest compression ratio examined (38:1) the correctness of the examiners' identification decisions is not affected in any statistically significant way, the study does not attempt to measure either the total effort necessary to reach these decisions or the examiners' confidence in those decisions.

Finally, application of lossy compression to latent imagery (for comparison to lossless exemplar imagery) shows promise as all cases examined appeared to be statistically indistinguishable from the lossless control case by subjective analysis and judgment, up to a compression rate of 5:1. Examination of error rates showed that lossy compressed latent imagery did not result in a measurable change in examiner error rates up to and including the maximum treatment of 38:1 when compared to a lossless exemplar. While it is premature to recommend lossy compression for latent fingerprint imagery given the many variables involved due to the uncontrolled nature of latent fingerprints, the data in this study does warrant further study on the matter, especially since 5 of the 6 test cases showed anecdotal evidence (not statistically significant) that a rate other than 1:1 lossless control case is optimal under both subjective and objective judgment criteria (see sections 4.3.1 and 4.4.1 respectively).

VERSION HISTORY

Date	Activity

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

TERMS AND DEFINITIONS

The abbreviations and acronyms of Table 1 are used in many parts of this document.

Table 1 - Abbreviations

NIST	National Institute of Standards and Technology
JPEG	Joint Photographic Experts Group – ISO/IEC committee developing standards for image compression
IAI	International Association for Identification
FBI	Federal Bureau of Investigation
SIVV	Spectral Image Validation/Verification Metric
NBIS	NIST Biometric Image Software
WSQ	The Wavelet Scalar Quantization algorithm for compression of fingerprint imagery
NGI	Next Generation Identification
IAFIS	Integrated Automated Fingerprint Identification System

ABSTRACT

This paper presents the findings of a study conducted to measure the impact of JPEG 2000 lossy compression on the comparison of 1000 ppi latent fingerprint imagery and 1000 ppi exemplar fingerprint imagery. Combinations of image pairs that vary by the compression rate applied to one of the images in the pair are observed and analyzed. The impact of lossy compression to both Galton and non-Galton based features of a fingerprint is measured by professional judgment of expert fingerprint examiners. The impact of compression is analyzed by quantifying multiple decisions relative to different levels of loss incurred during image compression. In addition to measuring the perceived visual impact of compression on the aforementioned features of the fingerprint, this paper also looks at the impact of lossy compression on the examiner's ability to correctly render their identification decisions.

KEYWORDS

1. Investigative Goals and Objectives

In 2009, NIST commenced an investigation on the use of JPEG 2000 [JPEG2K] for compressing 1000 ppi fingerprint imagery in partnership with the Federal Bureau of Investigation (FBI). While addressing this set of objectives, the investigators set out to examine the impact of JPEG 2000 compression on latent casework, which is the topic of this study. The specific goals and objectives are as follows:

1. **Identify an optimal exemplar compression rate with respect to latent casework:** Determine an effective and optimal compression target ratio for 1000 ppi exemplar imagery intended for comparison to lossless latent fingerprint imagery.
2. **Assess the impact of exemplar lossy compression on latent identification error rates:** Assess whether lossy compression of exemplar fingerprints results in any detectable increase in the identification error rates when these exemplars are compared to lossless latent fingerprint images.
3. **Explore the impact of latent fingerprint lossy compression on latent examiner subjective evaluation of image degradation:** Conduct an exploratory study on the application of lossy compression to latent fingerprint imagery and survey the impact of such compression on examiner assessment of subjective evaluation of degradation when compared to lossless exemplar fingerprints.
4. **Explore the impact of latent fingerprint lossy compression on latent examiner identification error rates:** Conduct an exploratory study on the application of lossy compression to latent fingerprint imagery and survey the impact of such compression on examiner identification error rates when comparing these images to lossless exemplar fingerprints.

1.1. Background

Criminal justice communities throughout the world exchange fingerprint data primarily in 8-bit gray-scale at 500 pixels per inch (ppi). The Wavelet Scalar Quantization (WSQ) fingerprint image compression algorithm is currently the standard algorithm for the compression of exemplar 500 ppi fingerprint imagery, and various lossless standard compression algorithms may be applied for latent 500 ppi fingerprint imagery. The WSQ standard defines a class of encoders and decoders with sufficient interoperability to ensure that images encoded by one compliant encoder can be decoded by any other compliant decoder.

As next generation fingerprint systems transition to 1000 ppi resolution images, practitioners will need to use JPEG 2000 for compression since WSQ was designed and optimized for 500 ppi only. For this reason, the FBI and NIST are exploring the behavior and limitations of JPEG 2000 on 1000ppi friction ridge imagery inclusive of rolled, flat and latent impressions.

A study conducted by the International Association for Identification (IAI) [FITZPATRICK] established 15:1 as the target compression ratio for the WSQ algorithm that would retain acceptable exemplar image fidelity for 500 ppi fingerprint imagery. The IAI study used the opinions of expert fingerprint examiners to measure image fidelity degradation due to lossy compression. Utilizing examiner opinion does not imply that automated fingerprint matcher performance is not an important criterion in a given biometric system, but if fingerprints are to be admissible as evidence in a court of law their ultimate utility lies in the expert examiner's opinion of the fidelity of those fingerprints. In 2011, NISTIR-7778 [NIST2] built upon the IAI study with expanded scope for 1000 ppi fingerprint imagery, but again looking at exemplar images resulting from controlled fingerprint capture only. NISTIR-7778 showed that a target compression rate of 10:1 may be optimal for exemplar images from controlled fingerprint capture. NISTIR-7778 did not however examine the impact of the various lossy compression rates on exemplar fingerprint imagery when used in latent fingerprint-matching case work which includes images typically from uncontrolled fingerprint capture. That is the focus of this study.

1.2. Key Drivers and Mandates

The FBI is modernizing its technology environment with the Next Generation Identification (NGI) initiative, part of which involves upgrading systems for the handling of fingerprints digitized at 1000 ppi. This is part of an overarching effort to improve fingerprint identification and verification tasks and meet its mandate.

NIST has an established expertise in evaluating biometric systems and standards, and has been assigned by the USA PATRIOT Act (Public Law 107-56) the responsibility for developing and certifying biometric technology standards. NIST has been supporting biometric standards and evaluation activities for over forty five years, starting with fingerprint analysis which began in 1965.

In support of the FBI, the National Institute of Standards and Technology (NIST) is conducting a set of studies that explore and compare image compression algorithm behavior at the 500 and 1000 ppi levels. The ultimate goal from these studies is to build normative compression guidance at 1000 ppi to enhance the ANSI/NIST ITL 1-2011 standard [AN29]. The ANSI/NIST standard serves as the basis for FBI's [EBTS] standard.

2.   Materials and Methods

In order to create the dataset of fingerprint images to be used in this experiment, selections were made from NIST Special Database 27 [SD27]. SD27 contains 291 latent fingerprint impressions along with known matches to rolled and flat exemplar fingerprint images also within SD27.

The image set resulting from this selection process contained various image impression types as described in Table 2 from a total of 175 unique individuals (25 subjects have contributed more than one unique latent fingerprint impression to SD27) as described in Appendix A. Also, as noted in Table 2, image pairs were selected from both match samples (where the two images were from the same finger/same person) as well as non-match samples of corresponding fingers of the same Galton [GALTON] three-pattern classification (the same pattern of either Arch, Loop or Whorl, but from different individuals and therefore non-matching).

Table 2 - Comparison Cases for This Study

Case Number	Probe Image (Lossless)	Gallery Image (Lossy Compressed)	Mated Pair	Key Operational Cases of Interest	Case Count
1	Flat	Latent	Yes		200
2	Flat	Latent	No		200
3	Latent	Flat	Yes	✓	200
4	Latent	Flat	No	✓	200
5	Rolled	Latent	Yes		200
6	Rolled	Latent	No		200
7	Latent	Rolled	Yes	✓	200
8	Latent	Rolled	No	✓	200
9	Latent	Latent	Yes		200
10	Latent	Latent	No		200
					Total: 2,000

Each pairing from the 10 cases in Table 2 consists of a probe image and an exemplar image. The probe image is a latent or exemplar fingerprint image with no lossy compression treatments applied in that image's lifecycle. The gallery image for each of the 10 cases described in Table 2 was generated by processing that image using each of the 14 compression treatments (ratios) described in Table 3, and then decompressing the image. The compression treatments utilized in this study were lossy in all cases except that for the 1:1 ratio (lossless / control case) which was not processed by any compression algorithm. After going through the process of being compressed and decompressed, these processed images have now passed through the typical stages in the lifecycle of images in a normal operational environment that occur just prior to processing/enrollment. Each image pair for every case in Table 2 is then presented to the examiners as described in Section 2.2. Further details on the selection of images for this study are described in Appendix A.

It should also be noted that images may be decompressed many subsequent times following the initial compression pass, but the decompression process has no effect on the compressed stream and therefore is not a topic of study. The images were compressed only once, thereby limiting compression degradation to only that resulting from the first compression pass.

The compressed images resulting from processing, along with the control images, were then paired with the original image, creating a set of 28, 000 image pairs as shown in Table 3.

Table 3 - Pairing of Each Probe (Original Image) to its Exemplar at Increasing Levels of Compression

Pair Number	Probe	Exemplar
1	Original Image	1 to 1 (control, Original Image)
2	Original Image	2 to 1
3	Original Image	5 to 1
4	Original Image	7 to 1
5	Original Image	10 to 1
6	Original Image	12 to 1
7	Original Image	15 to 1
8	Original Image	17 to 1
9	Original Image	20 to 1
10	Original Image	22 to 1
11	Original Image	26 to 1
12	Original Image	30 to 1
13	Original Image	34 to 1
14	Original Image	38 to 1

Of the ten pairings in Table 2, two of them are of critical operational importance. These include the cases of matching a never-before-compressed latent image to a lossy-compressed flat exemplar fingerprint from the same person (case 3 from Table 2) and matching a never-before-compressed latent image to a lossy-compressed rolled exemplar fingerprint from the same person (case 7 from Table 2). Another case of operational interest is that of comparing a latent fingerprint to another latent from the same person, but this case is not specifically addressed in this study and is outside of the study's investigative goals and scope.

2.1. Compression Algorithm

The primary focus of this study is to measure the effects of the application of lossy compression using the JPEG 2000 algorithm [JPEG2K] to exemplar imagery intended for comparison with lossless latent imagery. JPEG 2000 is an image compression standard and coding system that was created by the Joint Photographic Experts Group committee (JPEG) in 2000 to improve on the original JPEG image compression standard's discrete cosine transform-based methodology [JPEG] by utilizing a wavelet-based methodology. This modification yielded increases in both data compression and subjective image quality. Moreover, JPEG 2000 provides additional flexibility in the creation and manipulation of the code-stream and is based on the same family of wavelets as WSQ which is currently the standard for fingerprint image compression at 500 ppi. The flexibility offered by JPEG 2000 as well as the greater availability of JPEG 2000 implementations, which are commodity products as opposed to the much more specialized WSQ implementations, make JPEG 2000 an appropriate candidate as the successor to WSQ.

The implementation of JPEG 2000 used in this experiment was Open JPEG's [OPENJPEG] reference implementation version 1.3. This reference implementation has been incorporated into the NIST Biometric Image Software (NBIS) public domain software distribution [NIST1].

For this experiment it was necessary to create a specially tailored compression approach to generate test images for all necessary compression ratios being examined. This tailoring strategy utilized several exploratory studies which established bounds and baselines for parameters such as the compression ratio, intermediate compression layers and decomposition levels. This tailoring yielded 13 sets of parameters, one set for each of the 13 compression ratios to be investigated. Each of the original images used was processed at each of the 13 compression levels, yielding 13 compressed images plus one control image that was not compressed. Thus each original image used in the study yielded 14 test images. The specially tailored compression approach was based on the guidance found in the informative "Profile for 1000 ppi Fingerprint Compression" [MTR1]. The guidance in [MTR1], which called for 6 decomposition levels, was adjusted to use 5 decomposition levels using data from the exploratory studies in creating the tailored parameter sets in Table 5. The guidance also called for 9 quality layers which was tailored accordingly to accommodate the top compression layers called for in this study. These target compression ratios and associated quality layers are listed in Table 4. Note that JPEG 2000 enables structuring of the code stream such that it may be progressively decompressed at any of a series of intermediate compression levels (quality layers) in addition to the final "target" compression level. This feature of JPEG 2000 is intended to allow for the display of lower fidelity versions of the image to suit, for example, lower resolution displays while adding negligibly to the size of the compressed data stream and having negligible effect on the image at the target compression level. Other specific configuration parameters used for the JPEG 2000 compression of images are provided in Table 5.

Table 4 - Quality Layers Used in Study

Target Compression Ratio	All Quality Layers Including Top Layer[2]
1 to 1	Control image, no processing was performed
2 to 1	2, 10, 15, 24, 34, 58, 86, 144, 214
5 to 1	5, 10, 15, 24, 34, 58, 86, 144, 214
7 to 1	7, 10, 15, 24, 34, 58, 86, 144, 214
10 to 1	10, 15, 24, 34, 58, 86, 144, 214, 324
12 to 1	12, 15, 24, 34, 58, 86, 144, 214, 324
15 to 1	15, 24, 34, 58, 86, 144, 214, 324, 540
17 to 1	17, 24, 34, 58, 86, 144, 214, 324, 540
20 to 1	20, 24, 34, 58, 86, 144, 214, 324, 540
22 to 1	22, 24, 34, 58, 86, 144, 214, 324, 540
26 to 1	26, 34, 58, 86, 144, 214, 324, 540, 980
30 to 1	30, 34, 58, 86, 144, 214, 324, 540, 980
34 to 1	34, 58, 86, 144, 214, 324, 540, 980, 1930
38 to 1	38, 58, 86, 144, 214, 324, 540, 980, 1930

Table 5 - Other JPEG 2000 Compressor Settings Used in Study

Compressor Configuration Setting	Description
-n 6	6 resolution levels (original + 5 levels of decomposition[3])
-p RPCL	Resolution-Position-Component-Layer (RPCL) progression order
-b 64, 64	Code block size of 64x64
-r [rate values from Table 4]	Specifies the target top-layer rate, plus other quality layers
-d 0,0	Image origin offset
-I	Use irreversible compression (lossy)
-S 1,1	Use subsampling factor of 1,1

---

[2] Note that the higher target compression ratios and intermediate quality layers indicated in the guidance were found to cause malfunction of the Open JPEG 2000 v.1.3 encoder. Partially influenced by the image size, intermediate layers exceeding 324:1 to 540:1 caused the algorithm to return a minimally compressed image rather than the expected ratio. Consequently, problem images were re-run using more modest intermediate layer specifications. The intermediate levels were found to have no effect on the output image at the target compression ratio. This issue has been corrected in v.1.4 of the Open JPEG 2000 encoder.

[3] The Open JPEG 2000 codec sets the number of decomposition levels to one less than the value specified by this command line parameter. Hence, -n 6 yields 5 decomposition levels.

2.2. Methodology

Each image pair from Table 2 was shown to exactly three latent fingerprint examiners. Each examiner was first asked to determine if the image pair being displayed constituted a matched pair from the same individual (referred to throughout this document as the "Identification decision" or "Non-Identification decision"). Their responses in determining the identity of the pair being presented to them can be one of three choices:

- The presented image pair is from the same individual ([positive] Identification decision);
- The presented image pair does not appear to be from the same individual (Non-Identification decision);
- A decision of identity cannot be made (Inconclusive).

The individual decisions above were verified against the answer key and the results were recorded as a series of codes noted in Table 6 below.

Table 6 - Observation Codes for Decision of Identity

Observation Numeric Code	Description
1	A correct decision of identity or non-identity was made.
2	An inconclusive decision was made.
3	An incorrect decision of identity or non-identity was made.

For each pair of images examined, a match decision was provided by 3 different examiners. The three decisions were combined into a single decision triplet and sorted by digits. For example, the decision triplets "1,2,1", "1,1,2" and "2,1,1" were each sorted to decision code "112". The identity decision triplets are provided in Table 7 below.

Table 7 - Sorted Identification Decision Triplets and Justifications

Responses from Observers	Justification
111	These cases represent a majority-correct decision by the 3 examiners.
112	
113	
121	These cases represent a majority-inconclusive decision by the examiners. Split decision cases (1, 2, 3) are also considered inconclusive in this study.
122	
123	
221	
222	
223	
331	These cases represent a majority-incorrect decision by the examiners.
332	
333	

Subsequent to their decision of identity for the image pair, each examiner was then asked to evaluate that image pair for fidelity loss. To aid in analysis and quantification of fidelity loss, the examiner's evaluation was collected by utilizing a Likert response scale [LIKERT]. The responses that the examiners could make are provided in Table 8. Observation codes representing the response scale are ordered in ascending order indicating a progressively greater amount of degradation from 1 to 4. Furthermore, the features summarized in observation code 3 are among those typically used for forensic-level decisions while features summarized in observation code 4 are those used primarily by automated-matchers in rendering a match decision.

Table 8 - Observation Codes for Compression Degradation Observation

Observation Numeric Code	Description
1	No apparent image quality degradation and the quality of Level 2 [4] and Level 3 detail in either image should not cause any difficulty in reaching a conclusive decision of identification or exclusion.
2	A noticeable degradation in the quality of Level 2 or Level 3 detail in either image, but not enough to have a negative impact on reaching a conclusive decision of identification or exclusion, though the amount of time to reach a decision may increase.
3	Level 3 detail quality diminished in either image to the extent that a Level 3 identification is questionable or not possible, and/or is significantly more difficult.
4	Level 2 detail quality diminished in either image to the extent that a Level 2 identification becomes questionable or not possible, and/or is significantly more difficult.

The observation codes for degradation provided by each of the three examiners were combined into a triplet and sorted by digit. For example, the decision triplets "1,2,1", "1,1,2" and "2,1,1" each sort to decision code "112". The degradation decision triplets are provided in Table 9 below.

Table 9 - Sorted Observation Rating Triplets and Justifications

Sorted Responses from Observers	Justification
111 112 113 114	These cases represent either a unanimous or majority ruling of observation code 1 from Table 8 indicating no apparent image degradation.
122 123 124 222 223 224	These cases represent either a unanimous or majority ruling of observation code 2 from Table 8. This case also includes the special split-decision cases of **1, 2, 3** and **1, 2, 4**. These cases are considered borderline cases but are assigned to this bin as they are biased towards an acceptable image rating by two examiners noting little or no observable loss.
134 234 133 233 333 334	These cases represent either a unanimous or majority ruling of observation code 3 from Table 8. This case also includes the special split-decision cases of 1, 3, 4 and 2, 3, 4. These cases are considered biased-towards, and indicative of level-3 detail loss.
144 244 344 444	These cases represent either a unanimous or majority ruling of observation code 4 from Table 8.

---

[4] The commonly accepted nomenclature defines Level 1 fingerprint details as the overall friction ridge pattern and flow, Level 2 detail as classic Galton features [GALTON] like minutiae points, and Level 3 as pores, creases, line shapes, incipient ridges and other non Level 1 or 2 features [JAIN].

Examiner responses were recorded by a custom test apparatus consisting of a commodity computer and custom testing software, the Fingerprint Image eXaminer's Tool (FIXT), developed specifically for this study. The examiners were not provided any time limits on their responses.

The image pairs were queued on each examiner's workstation and their presentation order was shuffled randomly on each of the three workstations.

The examiners were provided the basic ability to independently reposition (pan), rotate, invert and zoom in and out of each of the two images from the pair being examined. This provided them with the basic tools that they typically employ in their standard operating environments. While they were provided with basic tools, more advanced assistive technologies normally available to some examiners, such as on-screen feature marking or pixel intensity adjustments were not provided to them to maintain experimental control.

The image pairs were presented on a split-screen to the examiners in randomized orders (see Figure 1), and randomized placement (left/right split screen placement) to mitigate potential order effects or positional bias. Scientists overseeing the tests were blind to the placement order of the images, as well as to the compression level of image pairs. These factors were tracked by the test apparatus.

Figure 1 - Split-Screen Presentation of Image Pairs

Each image pair was guaranteed to be observed by 3 different examiners over the course of the study without repeating. Due to the physical limitations of even the best modern monitors, it is impossible to show a 1000 ppi image without zooming or interpolation (see Appendix B). The software apparatus enabled the examiner to view images at approximately 10x to 50x of the original size.

Once the examiner made a decision for a given pair of images, that pair was marked as complete. The examiner was allowed to return to a completed pair and re-examine that image pair without penalty. The examiner was also allowed to jump to any image pair in the queue regardless of that pair's position in the examiner's queue.

Examiners were provided basic verbal instructions and a demonstration on how to use their workstations. They were allowed a brief practice session using image pairings that were not part of the study to gain familiarity with the procedure, scoring, and workstation controls. Examiners were allowed to freely ask questions for clarification on their workstations or tasking. The examiners were located in the same room and were allowed to interact freely as they do in their normal professional practice. Finally, the examiners were advised that once they've selected one of the three workstations on which to process images, they continue to use that workstation exclusively. This was done in order to eliminate the possibility of an examiner processing the same image pair more than once.

2.3.   Participants

This study utilized 20 paid professional latent fingerprint examiners to look at fingerprint image pairs in their professional capacity and to render their professional judgment much as they do in their normal professional activities.

There was no attempt to evaluate an examiner's level of proficiency for this study either prior, during or after the study. There was also no attempt to identify or maintain the identity of the individual examiners utilized in this study.

The examiners had anywhere from two years to over forty years of experience in fingerprint examination.

A pre-requisite for examiners being selected to participate in this study was that they be trained *latent-print-examiners (LPE's)* as the experimental cases in this study all included latent imagery. Of the 20 examiners participating in this study, 18 had earned the IAI's latent examiner certification (and were referred to as *certified latent print examiners*, or *CLPE's*) with the remaining two being non-certified latent examiners.

The examiners were recruited from various Federal, State, local and commercial entities and were permanently based in 11 states: California, Colorado, Florida, Illinois, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Texas, and Virginia.

3. Analysis

3.1. Normality of Degradation Score

In order to apply hypothesis tests as described in the next section, we examined the degree to which the distributions of the degradation score approximate a standard normal distribution. We first tried applying the normality assumption on the basis of inspection of histograms with overlain normal fitted functions and normal probability plots. A typical example is displayed in Figure 2 and Figure 3 below. The data histogram in Figure 2 shows informally that the sample data is not a good fit to a normal probability curve and appears bimodal in distribution. The empirical distribution of the histogram should be bell-shaped and resemble the normal distribution. This might be difficult to see if the sample is small. In this case one might proceed by regressing the data against the quantiles of a normal distribution with the same mean and variance as the sample. Lack of fit to the regression line suggests a departure from normality. The normal probability plot [CHAMBERS] provides a nonparametric means by which to compare quantiles of two distributions. Distributions may be taken as normal if a quantile plot overlays that of the standard normal distribution (a straight line in this type of plot) without major departures from linearity. Application of Shapiro-Wilks normality test [SHAPIRO] to provide a quantitative basis for normality confirmed that the data does not follow a normal distribution (Shapiro-Wilk $p<0.0001$). Hence, hypothesis testing employed non-parametric analysis methods.
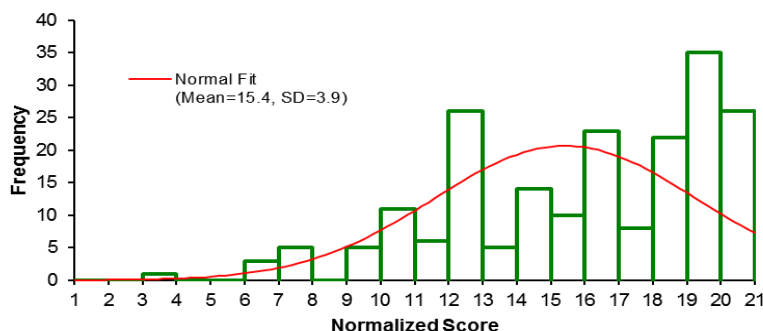
Figure 2 - Histogram typical of normalized degradation score with the normal distribution function fitted to data.
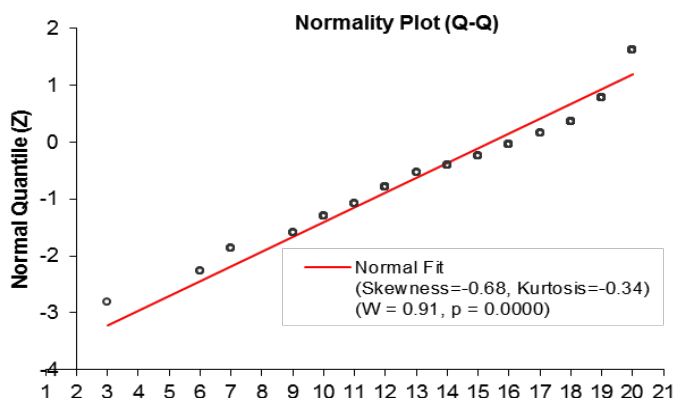
Figure 3 - Normal probability plot typical of degradation score data. Normality is indicated by linearity.

3.2.   Analysis Process

As described in section 2.1 this study examines 13 treatments and one control case. In stage one of the analysis process, Friedman's test was applied to the score results from a given experimental case and the treatments were sorted according to mean rank. The Friedman test [CONOVER] is a non-parametric test for analyzing randomized block designs where every treatment is replicated the same number of times and is an extension of the sign test for when there may be two or more treatments of interest. The Friedman test assumes that there are two or more experimental treatments (k). The observations are arranged in b blocks as in the example shown in Table 10.

Table 10 - Block Design Example

Block (i)	Treatment (j)			
	1	**2**	**...**	**k**
1	X_{11}	X_{12}	...	X_{1k}
2	X_{21}	X_{22}	...	X_{2k}
3	X_{31}	X_{32}	...	X_{3k}
...	...	...	...	...
b	X_{b1}	X_{b2}	...	X_{bk}

Let $R(X_{ij})$ be the rank assigned to X_{ij} within block i, and in the case of ties the average rank is used.
The ranks are summed to obtain

$$R_j = \sum_{i=1}^{b} R\left(X_{i,j}\right)$$ (1)

for j = 1, 2, ..., k.

Then the Friedman test is

H$_0$:   Null hypothesis, the treatment effects have identical effects ($\mu_0 = \mu_1$)
H$_a$:   Alternative hypothesis, at least one treatment is different from at least one other treatment ($\mu_0 \neq \mu_1$)

where μ_j is the mean rank of treatment j, j = 1 ... k.

Test Statistic:

$$T_1 = \frac{12}{bk(k+1)} \sum_{i=1}^{k} \left(R_i - \frac{b(k+1)}{2} \right)^2$$ (2)

If there are ties, then

$$T_1 = \frac{(k-1) \sum_{i=1}^{k} \left(R_i - \frac{b(k+1)}{2} \right)^2}{A_1 - C_1}$$ (3)

where

$$A_1 = \sum_{i=1}^{b} \sum_{j=1}^{k} \left(R\left(X_{ij}\right) \right)^2$$ (4)

$$C_1 = \frac{bk(k+1)^2}{4}$$ (5)

Note that Conover recommends the statistic

$$T_2 = \frac{(b-1)T_1}{b(k-1) - T_1}$$ (6)

since it has a more accurate approximate distribution. The T_2 statistic is the two-way analysis of variance statistic computed on the ranks $R(X_{ij})$.

The significance level, or the accepted probability of incorrectly rejecting the null hypothesis (i.e., making a Type I error) for the Friedman's test is set at 0.05.

For the T_2 test statistic, the critical region for rejection of the null hypothesis is

$$T_2 > F_{(\alpha, k-1, (b-1)(k-1))} \qquad (7)$$

where $F_{(\alpha, k-1, (b-1)(k-1))}$ is the percent point function of the F distribution.

For the T_1 test statistic, the critical region is given as

$$T_1 = X^2_{(\alpha, k-1)} \qquad (8)$$

where X^2 is the percent point function of the chi-square distribution.

The T_1 approximation is sometimes poor, so the T_2 approximation is typically preferred.

Friedman's analysis provides a mechanism to order the treatments. The treatments are sorted by their mean rank and then provided an [ordinal] rank number ranging from 1 (best) to 14 (worst).

Friedman's test can be used to determine of one of the treatments differs from at least one other treatment. If the Friedman's test yields a probability greater than or equal to the selected Type I error threshold, the alpha rate, we accept the null hypothesis that there are no differences among treatments. If the Friedman's test results in a probability value less than the selected alpha rate, we reject the null hypothesis and accept the alternative hypothesis that at least one of the treatments groups has been sampled from a different population. In order to determine which treatments are different, we perform a post hoc analysis of the pairwise differences between the treatments.

Commonly, Friedman's test is followed by *post hoc* analysis involving multiple pairwise comparisons to determine which of the treatments differ from one another at levels of statistical significance. For purposes of the present study, however, the multiple pairwise comparison procedure was used only to identify which of the cluster of alternative treatment near the top of the mean rank score list differed significantly from the treatment occupying the top position.

The Wilcoxon Signed Rank Test [WILCOXON], [CONOVER], [HOLLANDER] was used for the limited number of pairwise comparisons. The Wilcoxon Signed Rank Test is the non-parametric analog to the pairwise t-test. It examines the pairwise (intra-block) differences between measurements, $|Y_i-X_i|$, to test the null hypothesis that the expected value of the pairwise difference in measurements, $E(D)$, is zero for the two-tailed test.

$$H_0: \quad E(D) = 0 \quad (i.e., E(Y_i) = E(X_i))$$
$$H_1: \quad E(D) \neq 0$$

Given the ordering of the compression treatments as indicated by the mean rank scores, the Wilcoxon test was applied to examiner observation data for pairs of treatments as ordered by mean rank scores so as to successively compare pairwise measurements of the treatment deemed the best with those deemed next best, and so on down the list until a significant difference was found as denoted by the p value calculated by the Wilcoxon test. Thus, we were able to assess if a treatment deemed best by the mean rank criterion was of statistically significant difference from potential contenders. Moreover, the ordering of the treatments with successive pairwise testing obviated the need to test every treatment against every other treatment. The importance of limiting the number of comparisons is important will become apparent from the discussion which follows.

Use of *post hoc* analysis requires the use of some measure to control the experiment-wise Type I error rate, i.e., the probability of incorrectly rejecting the null hypothesis. The philosophical underpinnings are explained in [CURRAN-EVERETT], but essentially the idea is that where an experiment consists of multiple decisions, each carrying a probability of error, the likelihood of making an incorrect decision increases with the number of decisions. Accordingly, for an experiment involving multiple planned, dependent (non-orthogonal) decisions to accept or reject the null hypothesis, the accepted error probability for incorrectly rejecting the null hypothesis must be apportioned among the various decisions

such that the overall error probability does not exceed the acceptable error rate. Various methods for controlling this "family-wise" Type I error are described in [CONOVER] and [SOKAL].

We employ the Bonferonni method to control the experiment-wise error rate, α. For each of a series of k pairwise comparisons, the adjusted significance level, α', to be used for each comparison in the experiment, may be computed as

$$\alpha'_k = \frac{\alpha}{k} \tag{9}$$

where k = the number of pairwise comparisons. Hence, given a set of compression treatments ordered by mean rank score, comparison of each of several contending treatments may be compared to that in the top position via the Wilcoxon test which tests the probability that the two treatments yield equivalent results. For each comparison, we adjust the threshold for acceptance of the null hypothesis according to equation (14). We are not interested in all pairwise comparisons, but at most only in comparisons involving the top ranking treatment of each of other treatments of a group. Thus, we proceed down the list of treatments ordered by mean rank to test the difference between the top ranking treatment and others until we find a statistically significant difference. We group these treatments as statistically not differentiable from each other. The treatment that was found to be statistically different from the top-ranked of the grouping just formed becomes the top-ranked of the next group to which we compare with treatments below it in the list of mean ranks. Again, we form a group of treatments when we encounter a significant difference. We proceed similarly to the end of the list. At the end of the procedure we have a set of group members that are statistically similar to their top ranked treatment.

As mentioned above, we are not interested in all possible pairwise comparisons, but rather at most only comparisons of treatments to that having the top position when ordered by mean rank. Thus, we have at most 13 pairwise comparisons, i.e. k =13 and conservatively use the adjusted alpha value $\alpha' = 0.05 / 13 = 0.003846$ as the threshold for rejection of the null hypothesis in each pairwise test.

4. Results

4.1. Investigative Goal 1: Identify an optimal exemplar compression rate with respect to latent casework

In a decision system operating on signal data that has been processed by one or more signal processing treatments, the signal treatment generating the most favorable outcomes using a predetermined set of measurement criteria is identified as the optimal signal treatment. In the case of using lossy compression treatments on digital fingerprints, the compression treatment generating a fingerprint image with the highest level of fidelity to the original (non-lossy processed) fingerprint image is favored anecdotally and almost universally accepted as the best[5]. For the scope of this investigation, the optimal compression treatment is defined as the treatment yielding the best subjective ratings by fingerprint examiners over the other treatments. If more than one treatment is statistically indistinguishable from the best treatment (or clustered with) in terms of performance, the treatment with the highest achieved compression ratio within the cluster is deemed the optimal treatment in that it yields fidelity as good as the top rated treatment while at the same time achieving the best compression performance within the cluster of treatments.

4.1.1. Investigative Analysis 1

Data collection from the examiners yielded a series of 200 sorted examiner response triplets for each of the 14 treatments applied to four comparisons of interest for this investigative goal (cases 3, 4, 7 and 8 from Table 2). Friedman's test was applied to data from examiner observations of compression degradation in the case of lossy compressed exemplars to lossless latent fingerprints, for mated and non-mated image pairings of Rolled exemplars (Table 11) as well as Flat exemplars (Table 12) and the treatments ordered by mean rank. Friedman's test, significant with $p < 0.05$, indicated at least two treatment populations being statistically different from each other therefore warranting post-hoc analysis. The ordered list of treatments then underwent post-hoc analysis using Wilcoxon Signed Rank Test in a pairwise fashion as described in section 3 using the adjusted alpha level 0.0038.

Table 11 - Subjective Comparison of Lossy Compressed Rolled Exemplar Fingerprints vs. Lossless Latent Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank	Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
5:1	1	1217.0	6.09	NA	NA			A	5:1	1	1252.0	6.26	NA	NA			A
7:1	2	1288.0	6.44	1	2	0.1786		A	1:1	2	1295.5	6.48	1	2	0.6543		A
2:1	3	1312.5	6.56	1	3	0.1211		A	12:1	3	1355.0	6.78	1	3	0.1318		A
1:1	4	1313.0	6.57	1	4	0.1460		A	2:1	4	1359.0	6.80	1	4	0.0848		A
10:1	5	1350.5	6.75	1	5	0.0197		A	7:1	5	1447.0	7.24	1	5	0.0722		A
12:1	6	1417.0	7.09	1	6	0.0072		A	15:1	6	1465.0	7.33	1	6	0.0063		A
17:1	7	1417.5	7.09	1	7	0.0074		A	17:1	7	1483.0	7.42	1	7	0.0034	Yes	B
20:1	8	1489.5	7.45	1	8	0.0004	Yes	B	10:1	8	1484.5	7.42	7	8	0.7651		B
15:1	9	1516.5	7.58	8	9	0.8460		B	20:1	9	1522.5	7.61	7	9	0.6748		B
30:1	10	1680.0	8.40	8	10	0.0199		B	22:1	10	1580.5	7.90	7	10	0.1904		B
22:1	11	1711.0	8.56	8	11	0.0752		B	34:1	11	1630.5	8.15	7	11	0.0321		B
38:1	12	1754.5	8.77	8	12	0.0020	Yes	C	30:1	12	1639.0	8.20	7	12	0.0569		B
34:1	13	1763.0	8.82	12	13	0.6611		C	26:1	13	1724.0	8.62	7	13	0.0071		B
26:1	14	1770.0	8.85	12	14	0.8159		C	38:1	14	1762.5	8.81	7	14	0.0003	Yes	C
		Mated Exemplar (Friedman's p <0.0001)									Non-Mated Exemplar (Friedman's p <0.0001)							

---

[5] It has been shown the sometimes a slight amount of signal noise reduction applied to a noisy image yields an image that subjectively is favored by human examiners over its lossless counterpart.   This phenomenon is discussed further in 4.3.2.

Table 12 - Subjective Comparison of Lossy Compressed Flat Exemplar Fingerprints vs. Lossless Latent Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
5:1	1	1277.0	6.39	NA	NA			A
7:1	2	1318.0	6.59	1	2	0.7542		A
2:1	3	1342.0	6.71	1	3	0.5327		A
10:1	4	1411.5	7.06	1	4	0.0629		A
1:1	5	1437.0	7.19	1	5	0.0769		A
12:1	6	1475.5	7.38	1	6	0.0106		A
15:1	7	1493.0	7.47	1	7	0.0107		A
22:1	8	1521.5	7.61	1	8	0.0009	Yes	B
17:1	9	1523.0	7.62	8	9	0.7622		B
20:1	10	1524.5	7.62	8	10	0.7593		B
30:1	11	1587.5	7.94	8	11	0.2540		B
26:1	12	1653.0	8.27	8	12	0.0993		B
34:1	13	1675.0	8.38	8	13	0.1070		B
38:1	14	1761.5	8.81	8	14	0.0095		B

Mated Exemplar (Friedman's p <0.0001)

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
2:1	1	1299.5	6.50	NA	NA			A
5:1	2	1314.0	6.57	1	2	0.8517		A
1:1	3	1366.5	6.83	1	3	0.6892		A
7:1	4	1400.0	7.00	1	4	0.0878		A
10:1	5	1461.5	7.31	1	5	0.0828		A
17:1	6	1467.0	7.34	1	6	0.0143		A
12:1	7	1500.0	7.50	1	7	0.0414		A
22:1	8	1506.0	7.53	1	8	0.0184		A
15:1	9	1511.0	7.56	1	9	0.0021	Yes	B
30:1	10	1556.0	7.78	9	10	0.8726		B
20:1	11	1562.0	7.81	9	11	0.7912		B
34:1	12	1676.0	8.38	9	12	0.0874		B
26:1	13	1677.5	8.39	9	13	0.3278		B
38:1	14	1703.0	8.52	9	14	0.1159		B

Non-Mated Exemplar (Friedman's p <0.0001)

4.1.2. Investigative Result 1

Data from Table 11 and Table 12 above summarize that for each of the four experimental cases, the results of the experiment exhibit clusters of treatments where one or more treatments are indistinguishable from each other. These clusters are labeled as A, B, C, etc. In all four experimental cases (3, 4, 7 and 8 from Table 2), Cluster-A represents the best operational outcome and contains the 1:1 control case. What this indicates is that there is a clustering of several lossy compression treatments where the end result of compression is visually indistinguishable from the 1:1 control case and clustered with the 1:1 control case as assessed by human examiners.

Table 13 shows the clusters for the four experimental cases (3, 4, 7 and 8 from Table 2). The highest compression ratio (treatment) where cluster-A (best) overlaps across all four experimental cases is 12:1 and is highlighted with a dashed box in Table 13. This indicates that at 12:1 lossy compression, examiner assessment of lossy compression degradation is indistinguishable from the lossless case (1:1) for Rolled and Flat fingerprint impressions, in both mated and non-mated cases. For the scope of this investigation, we identify the optimal exemplar compression rate as the highest possible compression treatment where examiner's subjective assessment of degradation for all cases was indistinguishable from the lossless case, or 12:1 as the optimal case.

Table 13 - Treatment Clustering For Lossy Compressed Exemplar Fingerprints vs. Lossless Latent Fingerprints

Treatment	Lossy Rolled To Lossless Latent, Mated (case 7 from Table 2)	Lossy Rolled To Lossless Latent, Non-Mated (case 8 from Table 2)	Lossy Flat To Lossless Latent, Mated (case 3 from Table 2)	Lossy Flat To Lossless Latent, Non-Mated (case 4 from Table 2)
1:1	A	A	A	A
2:1	A	A	A	A
5:1	A	A	A	A
7:1	A	A	A	A
10:1	A	B	A	A
12:1	A	A	A	A
15:1	B	A	A	B
17:1	A	B	B	A
20:1	B	B	B	B
22:1	B	B	B	A
26:1	B	B	B	B
30:1	B	B	B	B
34:1	C	B	B	B
38:1	C	C	B	B

4.2. Investigative Goal 2: Assess impact of lossy exemplar compression on latent identification error rates

Determination of fingerprint image quality degradation can be subjective in nature where signs of quality loss may be perceived differently from one examiner to the next. A more objective measure of quality for the purposes of fingerprint identification is measuring the rate of success by the examiners in making their identification decision which is the goal of this investigation.

4.2.1. Investigative Analysis 2

Data collection from the examiners yielded a series of 200 examiner identification response triplets based on the examiner identification decisions for image pairings in each of the 14 treatments applied to four comparisons of interest for this investigation (cases 3, 4, 7 and 8 from Table 2). Friedman's test was applied to data from examiner identification decision scores in the cases of lossy compressed exemplars to lossless latent fingerprints, for mated and non-mated image pairings of Rolled exemplars (see Table 14) as well as Flat exemplars (see Table 15) and the treatments ordered by mean rank. Of the four cases, Friedman's test indicated that of only one warranted post-hoc testing (mated Rolled exemplars). The ordered list of treatments for mated Rolled exemplars underwent post-hoc analysis using Wilcoxon Signed Rank Test in a pairwise fashion as described in section 3.

Table 14 - Objective Comparison of Lossy Compressed Rolled Exemplar Fingerprints vs. Lossless Latent Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank	Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
22:1	1	1415.0	7.08	NA	NA			A	1:1	1	1467.5	7.34	NA	NA			A
7:1	2	1427.5	7.14	1	2	0.8924		A	20:1	2	1471.0	7.36	1	2	-		A
30:1	3	1434.5	7.17	1	3	0.9550		A	30:1	3	1476.0	7.38	1	3	-		A
10:1	4	1440.0	7.20	1	4	0.9267		A	2:1	4	1485.0	7.43	1	4	-		A
5:1	5	1461.5	7.31	1	5	0.8984		A	34:1	5	1492.0	7.46	1	5	-		A
15:1	6	1478.0	7.39	1	6	0.3865		A	17:1	6	1493.0	7.47	1	6	-		A
20:1	7	1499.5	7.50	1	7	0.4914		A	12:1	7	1497.0	7.49	1	7	-		A
12:1	8	1500.0	7.50	1	8	0.5391		A	5:1	8	1499.5	7.50	1	8	-		A
34:1	9	1513.5	7.57	1	9	0.1002		A	26:1	9	1504.0	7.52	1	9	-		A
26:1	10	1519.0	7.60	1	10	0.1471		A	7:1	10	1508.0	7.54	1	10	-		A
1:1	11	1526.0	7.63	1	11	0.1535		A	10:1	11	1515.0	7.58	1	11	-		A
17:1	12	1532.5	7.66	1	12	0.0930		A	38:1	12	1522.5	7.61	1	12	-		A
2:1	13	1594.5	7.97	1	13	0.0041		A	15:1	13	1527.5	7.64	1	13	-		A
38:1	14	1658.5	8.29	1	14	0.0016	Yes	B	22:1	14	1542.0	7.71	1	14	-		A
Mated Exemplar (Friedman's p =0.0297)									Non-Mated Exemplar (Friedman's p = 0.9538)								

Table 15 - Objective Comparison of Lossy Compressed Flat Exemplar Fingerprints vs. Lossless Latent Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank	Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_1 (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
5:1	1	1440.5	7.20	NA	NA			A	5:1	1	1438.5	7.19	NA	NA			A
20:1	2	1441.0	7.21	1	2	-		A	22:1	2	1457.0	7.29	1	2	-		A
17:1	3	1466.0	7.33	1	3	-		A	38:1	3	1465.0	7.33	1	3	-		A
12:1	4	1469.5	7.35	1	4	-		A	30:1	4	1483.0	7.42	1	4	-		A
22:1	5	1475.0	7.38	1	5	-		A	1:1	5	1491.5	7.46	1	5	-		A
34:1	6	1477.5	7.39	1	6	-		A	12:1	6	1495.5	7.48	1	6	-		A
7:1	7	1482.5	7.41	1	7	-		A	15:1	7	1498.5	7.49	1	7	-		A
15:1	8	1496.0	7.48	1	8	-		A	17:1	8	1501.0	7.51	1	8	-		A
1:1	9	1513.5	7.57	1	9	-		A	34:1	9	1519.5	7.60	1	9	-		A
10:1	10	1527.5	7.64	1	10	-		A	7:1	10	1521.5	7.61	1	10	-		A
30:1	11	1529.0	7.65	1	11	-		A	2:1	11	1525.0	7.63	1	11	-		A
26:1	12	1533.0	7.67	1	12	-		A	26:1	12	1529.0	7.65	1	12	-		A
38:1	13	1551.5	7.76	1	13	-		A	10:1	13	1533.0	7.67	1	13	-		A
2:1	14	1597.5	7.99	1	14	-		A	20:1	14	1542.0	7.71	1	14	-		A
		Mated Exemplar (Friedman's p =0.6425)									Non-Mated Exemplar (Friedman's p = 0.4773)						

4.2.2. Investigative Result 2

Analysis of examiner's identification errors indicates that only one of the four cases examined for this investigative goal showed any detectable change in error rates (mated Rolled exemplars). This statistically significant change in error rates occurred at the 38:1 compression ratio, which was the highest compression ratio treatment examined in this study. While the examiners are able to detect and note significant quality loss and degradation in the images through the various treatments, exemplar compression rates as high as 34:1 did not result in a statistically significant change in their ability to conduct their identification tasks on lossless latent fingerprint imagery given the quality of the latent fingerprints used in this study.

4.3. Investigative Goal 3: Explore the impact of latent fingerprint lossy compression on latent examiner subjective evaluation of image degradation

Latent fingerprint imagery typically starts off at a point of diminished quality (vs. controlled capture exemplars) due to the uncontrolled nature of the collection/capture process. Therefore, any further degradation of quality may simply diminish what little information is already present and any further loss should be avoided (See Figure 4). Because of this, utilization of lossy compression for latent imagery has been advised against. Anecdotal evidence, however, suggests that lossy compression may actually introduce a certain amount of beneficial signal processing by introducing noise-filtering (loss pass signal filtering) that had been assumed to not be beneficial to the identification process. The experimental results in this section seek to explore this question and investigate the potential benefits of compressing latent imagery in a lossy fashion by comparing lossy-compressed latent images to lossless compressed flat and rolled exemplars in order to isolate the cost/benefit of lossy compression relative to the latent image only.

| Lossless/Original Latent Fingerprint Image | Lossy (250:1) Compressed Latent Fingerprint Image | Lossy (500:1) Compressed Latent Fingerprint Image | Lossy (1000:1) Compressed Latent Fingerprint Image |

Figure 4 - Example of Latent Fingerprint Fidelity Degradation Due to Lossy Compression using JPEG 2000

4.3.1. Investigative Analysis 3

Data collection from the examiners yielded a series of 200 examiner response triplets for each of the 14 treatments applied to the six comparisons of interest for this investigative goal (cases 1, 2, 5, 6, 9 and 10 from Table 2). Friedman's test was applied to data from examiner observation of compression degradation in the case of lossy compressed latent fingerprint images to mated and non-mated image pairings of lossless Rolled exemplars (Table 16), lossless Flat exemplars (Table 17) and lossless Latent fingerprints (Table 18), and the treatments ordered by mean rank. In each case, Friedman's test indicated at least two treatment populations being statistically different from each other therefore warranting post-hoc analysis. The ordered list of treatments then underwent post-hoc analysis using Wilcoxon Signed Rank Test in a pairwise fashion as described in section 3.

Table 16 - Subjective Comparison of Lossy Latent Fingerprints vs. Lossless Rolled Exemplar Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
1:1	1	1117.5	5.59	NA	NA			A
2:1	2	1121.5	5.61	1	2	0.9844		A
5:1	3	1239.0	6.20	1	3	0.1040		A
7:1	4	1342.0	6.71	1	4	0.0015	Yes	B
10:1	5	1437.0	7.19	4	5	0.4595		B
15:1	6	1471.0	7.36	4	6	0.1916		B
12:1	7	1488.0	7.44	4	7	0.0376		B
17:1	8	1555.5	7.78	4	8	0.0052		B
22:1	9	1559.0	7.80	4	9	0.0077		B
26:1	10	1584.5	7.92	4	10	0.0069		B
20:1	11	1655.0	8.28	4	11	0.0002	Yes	C
30:1	12	1691.5	8.46	11	12	0.3710		C
34:1	13	1817.5	9.09	11	13	0.0178		C
38:1	14	1921.0	9.61	11	14	<0.0001	Yes	D

Mated Exemplar (Friedman's p <0.0001, Adjusted α = 0.003846)

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
2:1	1	1202.0	6.01	NA	NA			A
5:1	2	1204.0	6.02	1	2	0.6136		A
1:1	3	1213.0	6.07	1	3	0.9171		A
7:1	4	1374.5	6.87	1	4	0.0180		A
10:1	5	1403.0	7.02	1	5	0.0019	Yes	B
15:1	6	1489.0	7.45	5	6	0.4251		B
12:1	7	1507.0	7.54	5	7	0.4164		B
20:1	8	1533.5	7.67	5	8	0.1722		B
17:1	9	1574.0	7.87	5	9	0.1253		B
26:1	10	1602.5	8.01	5	10	0.0398		B
22:1	11	1611.0	8.06	5	11	0.0556		B
30:1	12	1678.0	8.39	5	12	0.0015	Yes	C
38:1	13	1798.0	8.99	12	13	0.1077		C
34:1	14	1810.5	9.05	12	14	0.1040		C

Non-Mated Exemplar (Friedman's p <0.0001, Adjusted α = 0.003846)

Table 17 - Subjective Comparison of Lossy Latent Fingerprints vs. Lossless Flat Exemplar Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
2:1	1	1070.0	5.35	NA	NA			A
5:1	2	1147.0	5.74	1	2	0.1716		A
1:1	3	1241.0	6.21	1	3	0.0059		A
7:1	4	1324.0	6.62	1	4	<0.0001	Yes	B
10:1	5	1423.5	7.12	4	5	0.1562		B
12:1	6	1459.0	7.30	5	6	0.0954		B
15:1	7	1491.0	7.46	5	7	0.0192		B
20:1	8	1532.5	7.66	5	8	0.0090		B
22:1	9	1600.5	8.00	5	9	0.0005	Yes	C
17:1	10	1637.0	8.19	9	10	0.6181		C
26:1	11	1674.5	8.37	9	11	0.4690		C
30:1	12	1762.5	8.81	9	12	0.0868		C
34:1	13	1774.5	8.87	9	13	0.0490		C
38:1	14	1863.0	9.32	9	14	0.0024	Yes	D

Mated Exemplar (Friedman's p <0.0001, Adjusted α = 0.003846)

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
2:1	1	1173.0	5.87	NA	NA			A
1:1	2	1230.5	6.15	1	2	0.7616		A
5:1	3	1252.5	6.26	1	3	0.7877		A
7:1	4	1252.5	6.26	1	4	0.3874		A
12:1	5	1404.0	7.02	1	5	0.0054		A
15:1	6	1489.5	7.45	1	6	0.0003	Yes	B
10:1	7	1536.0	7.68	6	7	0.5007		B
20:1	8	1538.5	7.69	6	8	0.5889		B
17:1	9	1549.0	7.75	6	9	0.2155		B
26:1	10	1578.5	7.89	6	10	0.3085		B
22:1	11	1625.5	8.13	6	11	0.0956		B
30:1	12	1770.0	8.85	6	12	<0.0001	Yes	C
34:1	13	1773.5	8.87	12	13	0.9693		C
38:1	14	1827.0	9.14	12	14	0.5779		C

Non-Mated Exemplar (Friedman's p <0.0001, Adjusted α = 0.003846)

Table 18 - Subjective Comparison of Lossy Latent Fingerprints vs. Lossless Latent Fingerprints

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
2:1	1	1046.5	5.23	NA	NA			A
1:1	2	1102.5	5.51	1	2	0.4124		A
5:1	3	1111.5	5.56	1	3	0.1312		A
7:1	4	1160.0	5.80	1	4	0.1468		A
10:1	5	1356.0	6.78	1	5	<0.0001	Yes	B
12:1	6	1426.5	7.13	5	6	0.6936		B
20:1	7	1490.5	7.45	5	7	0.0977		B
15:1	8	1532.0	7.66	5	8	0.0749		B
17:1	9	1574.5	7.87	5	9	0.0145		B
22:1	10	1690.0	8.45	5	10	<0.0001	Yes	C
26:1	11	1765.0	8.83	10	11	0.5069		C
34:1	12	1840.0	9.20	10	12	0.0235		C
30:1	13	1872.5	9.36	10	13	0.0161		C
38:1	14	2032.5	10.16	10	14	<0.0001	Yes	D

Mated Pair (Friedman's p <0.0001, Adjusted α = 0.003846)

Treatment	Rank Number	Rank sum	Mean rank	μ_0 (Rank No.)	μ_i (Rank No.)	Wilcoxon p	$\mu_0 \neq \mu_i$	Clustered Rank
5:1	1	1224.5	6.12	NA	NA			A
1:1	2	1275.5	6.38	1	2	0.5794		A
7:1	3	1288.5	6.44	1	3	0.4735		A
2:1	4	1337.0	6.69	1	4	0.2399		A
10:1	5	1411.0	7.06	1	5	0.0225		A
17:1	6	1444.0	7.22	1	6	0.0183		A
15:1	7	1452.5	7.26	1	7	0.0173		A
12:1	8	1469.5	7.35	1	8	0.0024	Yes	B
20:1	9	1525.5	7.63	8	9	0.2709		B
26:1	10	1610.5	8.05	8	10	0.1007		B
22:1	11	1630.5	8.15	8	11	0.0558		B
30:1	12	1691.0	8.46	8	12	0.0119		B
34:1	13	1766.0	8.83	8	13	0.0010	Yes	C
38:1	14	1874.5	9.37	13	14	0.1297		C

Non-Mated Pair (Friedman's p <0.0001, Adjusted α = 0.003846)

4.3.2. Investigative Result 3

Data from Table 16, Table 17 and Table 18 summarize that for each of the six experimental cases (1, 2, 5, 6, 9 and 10 from Table 2). The results of the experiment exhibit clusters of treatments where one or more treatments are indistinguishable from each other. These clusters are labeled as A, B, C, etc. In all six experimental cases, Cluster-A represents the best operational outcome and contains the 1:1 control case. This indicates that there are several lossy compression treatments that when applied to latent fingerprints result in compressed latent fingerprint imagery that is visually indistinguishable from the 1:1 control case as assessed by human examiners.

From Table 19 below, the highest compression ratio (treatment) where cluster-A (best) overlaps across all six experimental cases is 5:1. This indicates that up to a compression ratio of 5:1, examiner assessment of lossy compression of latent fingerprint imagery degradation is indistinguishable from the lossless case (1:1) in both mated and non-mated cases. Perhaps more interesting is the fact that 2:1 is identified as the top treatment for 4 of the 6 experimental cases. While 2:1's top spot in 4 of the 6 cases is anecdotal given the statistical tests, this data suggests that utilization of lossy compression for latent fingerprint imagery may provide some benefit and warrants further study.

Table 19 - Treatment Clustering For Lossy Compressed Latent Fingerprint Imagery

Treatment	Lossy Latent To Lossless Rolled, Mated	Lossy Latent To Lossless Rolled, Non-Mated	Lossy Latent To Lossless Flat, Mated	Lossy Latent To Lossless Flat, Non-Mated	Lossy Latent To Lossless Latent, Mated	Lossy Latent To Lossless Latent, Non-Mated
1:1	A	A	A	A	A	A
2:1	A	A	A	A	A	A
5:1	A	A	A	A	A	A
7:1	B	A	B	A	A	A
10:1	B	B	B	B	B	A
12:1	B	B	B	A	B	B
15:1	B	B	B	B	B	A
17:1	B	B	C	B	B	A
20:1	C	B	B	B	B	B
22:1	B	B	C	B	C	B
26:1	B	B	C	B	C	B
30:1	C	C	C	C	C	B
34:1	C	C	C	C	C	C
38:1	D	C	D	C	D	C

4.4. Investigative Goal 4: Explore the impact of latent fingerprint lossy compression on latent examiner identification error rates

Another concern that has been cited in calling for lossless compression of latent fingerprints has been that lossy compression of latent fingerprints may increase the rate of examiner error and diminish the ability of the examiner to conduct their identification decisions. We assess this by determining the extent to which examiner error rate increases as a function of increasing compression level applied to latent prints compared to lossless exemplars, where exemplars may be rolled, flat, or non-compressed latent fingerprints.

4.4.1. Investigative Analysis 4

Data collection from the examiners yielded a series of 200 identification response scores based on the examiner identification decisions for image pairings in each of the 14 treatments applied to six comparisons of interest for this investigative goal (cases 1, 2, 5, 6, 9 and 10 from Table 2). Friedman's test was applied to data from examiner identification decisions in the cases of lossy compressed latent fingerprints compared to various other impression types including Rolled, Flat, and latent gallery fingerprints, for both mated and non-mated image pairings and the treatments ordered by mean rank (Table 20). Of the 6 cases examined, none of the cases yielded any population differences according to the Friedman's test ($p > 0.05$) therefore not warranting any post-hoc analysis using Wilcoxon Signed Rank, and any observed benefits by one compression ratio over another is anecdotal.

Table 20 - Objective Comparison of Lossy Latent Fingerprints vs. Lossless Gallery Fingerprints

Treatment	Rank	Rank sum	Mean rank	Clustered Rank	Treatment	Rank	Rank sum	Mean rank	Clustered Rank	Treatment	Rank	Rank sum	Mean rank	Clustered Rank	Treatment	Rank	Rank sum	Mean rank	Clustered Rank	Treatment	Rank	Rank sum	Mean rank	Clustered Rank	Treatment	Rank	Rank sum	Mean rank	Clustered Rank
1:1	1	1386.5	6.93	A	30:1	1	1457.5	7.29	A	15:1	1	1449.0	7.25	A	17:1	1	1448.0	7.24	A	10:1	1	1483.0	7.42	A	15:1	1	1435.0	7.18	A
7:1	2	1420.0	7.10	A	15:1	2	1462.0	7.31	A	10:1	2	1459.5	7.30	A	22:1	2	1455.5	7.28	A	17:1	2	1490.0	7.45	A	7:1	2	1455.5	7.28	A
5:1	3	1445.5	7.23	A	1:1	3	1464.0	7.32	A	7:1	3	1468.0	7.34	A	20:1	3	1460.5	7.30	A	20:1	3	1490.0	7.45	A	12:1	3	1470.5	7.35	A
15:1	4	1465.0	7.33	A	26:1	4	1471.5	7.36	A	20:1	4	1474.0	7.37	A	7:1	4	1463.5	7.32	A	22:1	4	1491.0	7.46	A	17:1	4	1477.5	7.39	A
17:1	5	1475.0	7.38	A	7:1	5	1482.0	7.41	A	12:1	5	1480.0	7.40	A	15:1	5	1468.0	7.34	A	15:1	5	1496.5	7.48	A	22:1	5	1478.0	7.39	A
10:1	6	1482.0	7.41	A	12:1	6	1495.0	7.48	A	2:1	6	1487.5	7.44	A	5:1	6	1490.0	7.45	A	1:1	6	1497.0	7.49	A	20:1	6	1485.0	7.43	A
2:1	7	1495.0	7.48	A	2:1	7	1497.5	7.49	A	1:1	7	1499.5	7.50	A	2:1	7	1493.5	7.47	A	7:1	7	1497.0	7.49	A	34:1	7	1486.5	7.43	A
20:1	8	1504.5	7.52	A	5:1	8	1503.0	7.52	A	26:1	8	1503.0	7.55	A	26:1	8	1510.5	7.55	A	26:1	8	1497.0	7.49	A	5:1	8	1507.5	7.54	A
12:1	9	1517.5	7.59	A	20:1	9	1515.0	7.58	A	17:1	9	1515.5	7.58	A	1:1	9	1515.5	7.58	A	30:1	9	1504.5	7.52	A	2:1	9	1509.0	7.55	A
26:1	10	1526.0	7.63	A	22:1	10	1518.0	7.59	A	34:1	10	1518.5	7.59	A	30:1	10	1515.5	7.58	A	2:1	10	1505.0	7.53	A	1:1	10	1509.5	7.55	A
22:1	11	1549.0	7.75	A	38:1	11	1521.0	7.61	A	5:1	11	1526.5	7.63	A	10:1	11	1530.0	7.65	A	34:1	11	1509.5	7.55	A	26:1	11	1522.5	7.61	A
38:1	12	1551.0	7.76	A	10:1	12	1524.5	7.62	A	22:1	12	1529.0	7.65	A	12:1	12	1544.0	7.72	A	12:1	12	1510.0	7.55	A	30:1	12	1524.0	7.62	A
34:1	13	1580.5	7.90	A	17:1	13	1541.0	7.71	A	30:1	13	1529.5	7.65	A	34:1	13	1552.5	7.76	A	38:1	13	1510.5	7.55	A	38:1	13	1542.5	7.71	A
30:1	14	1602.0	8.01	A	34:1	14	1548.0	7.74	A	38:1	14	1560.0	7.80	A	38:1	14	1553.0	7.77	A	5:1	14	1519.0	7.60	A	10:1	14	1597.0	7.99	A
Lossless Rolled Exemplar vs. Lossy Latent										Lossless Flat Exemplar vs. Lossy Latent										Lossless Latent Exemplar vs. Lossy Latent									
Mated Exemplar					Non-Mated Exemplar					Mated Exemplar					Non-Mated Exemplar					Mated Pair					Non-Mated Pair				
Friedman's p = 0.1171					Friedman's p =0.6638					Friedman's p = 0.9602					Friedman's p =0.1707					Friedman's p = 0. 8294					Friedman's p =0.5318				

4.4.2. Investigative Result 4

Analysis of examiner's identification errors indicates that utilizing compressed latent fingerprint imagery in comparison to lossless exemplar images did not yield any statistically significant changes in error rates. This held true up to the maximum compression ratio examined in this study (38:1). Aggressive lossy compression up to 38:1 applied to latent imagery did not result in a statistically significant change in the examiner's ability to conduct their identification tasks against lossless exemplar fingerprint imagery as well as lossless latent fingerprint imagery.

5.   Conclusions

Lossy compression processes applied to any signal incurs a cost in terms of discarding information from the original signal. The challenge with lossy compression is to adjust the amount of incurred signal loss in such a way as to achieve the maximum amount of compression while maintaining full usability of that signal for its originally intended purpose.

Latent fingerprint casework is particularly challenging as the original signal (image) starts with a certain amount of inherent noise and degradation due to environmental and other factors related to the unconstrained deposition and collection of such impressions. Therefore it is difficult to establish criteria as to how any further signal loss will impact the ultimate utility of a latent fingerprint image. For the scope of this study, given evidence that ultimate fidelity lies in lossless compression, we set out to identify the highest compression using a lossy treatment that yields an image as good as the lossless using various experimental scenarios utilizing professional latent fingerprint examiners.

For the experimental scenarios that best representing typical operational casework where a fingerprint examiner is provided a lossless latent fingerprint to compare to a lossy compressed rolled or flat exemplar fingerprint and asked to make an identification decision (Investigative Goal 1), the maximum lossy compression ratio that can be applied to the exemplar and yield performance indistinguishable from that using a lossless rolled or flat exemplar is recommended to not exceed 12:1 (see Table 13).

Analysis of examiner's identification errors indicates that only one of the cases examined (mated Rolled exemplars) showed a statistically significant change in examiner error rates (Investigative Goal 2). This detectable change in error rate occurred at the upper range of compression examined in this study (38:1). While compression degradation did not manifest itself as a statistically measurable difference in error rates up to and including the compression rate of 34:1, it should be noted however that very high compression rates such as 34:1 may negatively impact the time the examiners need in conducting their tasks.

Examination of the potential application of lossy compression to latent imagery shows some utility where all test cases examined appeared to be statistically indistinguishable from the lossless case at up to a lossy compression rate of 5:1 (Investigative Goal 3). Examination of error rates showed that lossy compression of latent imagery did not yield a measurable change in examiner error rates up to and including the maximum treatment of 38:1 (Investigative Goal 4). While it is premature to recommend lossy compression for latent fingerprint imagery given the many variables involved due to the uncontrolled nature of latent fingerprints, the data in this study does suggest further study on the matter. This is especially true in light of the fact that of the 6 scenarios examined, 5 scenarios showed evidence (although not statistically significant) of a beneficial effect due to lossy compression being applied to the latent fingerprint image.

6.   Future Work

The application of lossy compression to latent fingerprint imagery warrants further work to identify any potential benefits from a moderate amount of lossy compression applied to latent imagery. This benefit would primarily be realized in latent operations and processing over limited bandwidth communications such as Mobile ID applications.

7.   References

Publications and Reports

CHAMBERS	Chambers, John; William Cleveland, Beat Kleiner, and Paul Tukey (1983). Graphical Methods for Data Analysis. Wadsworth
FITZPATRICK	Fitzpatrick, M. et al. 1994, "WSQ Compression / Decompression Algorithm Test Report", IAI Annual Conference.
GALTON	Galton, F. (2005). *Finger prints*. Mineola, NY: Dover Publications. (Original work published 1892)
JAIN	Jain, A., "Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 1, January 2007.
LIBERT	"A 1D Spectral Image Validation/Verification Metric for Fingerprints". Libert, J.M.; Grantham, J.; Orandi, S. August 19, 2009. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=903078. Retrieved 2011-01-12.
LIKERT	Likert, R. (1932). *A Technique for the Measurement of Attitudes,* Archives of Psychology 140, 55.
MTR1	"Profile for 1000 ppi Fingerprint Compression". Lepley, M.A. http://www.mitre.org/work/tech_papers/tech_papers_04/lepley_fingerprint/lepley_fingerprint.pdf. Retrieved 2011-01-11.
NIST1	"NIST Biometric Image Software". http://Fingerprint.nist.gov/NFIS/. Retrieved 2011-01-12.
NIST2	Orandi, S., Libert, J. M., Grantham, J. D., Ko, K., Wood, S.S., Wu, J. Effects of JPEG 2000 Image Compression on 1000 ppi Fingerprint Imagery, NIST Interagency Report 7778, National Institutes of Standards and Technology, Gaithersburg, MD. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=911122. Retrieved from on 09/01/2012.
OPENJPEG	"OpenJPEG library : an open source JPEG 2000 codec". http://www.openjpeg.org/index.php?menu=news. Retrieved 2011-01-12.
SD27	M.D. Garris & R.M. McCabe, "NIST Special Database 27: Fingerprint Minutiae from Latent and Matching Tenprint Images," NIST Technical Report NISTIR 6534 & CD-ROM, June 2000.
SHAPIRO	Shapiro, S. S.; Wilk, M. B. (1965). "An analysis of variance test for normality (complete samples)". *Biometrika***52** (3-4): 591–611. doi:10.1093/biomet/52.3-4.591. JSTOR 2333709MR205384.

Standards

|	
AN29	NIST Special Publication 500-290: American National Standard for Information Systems — *Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information* (ANSI/NIST ITL 1-2011). Approved November 9, 2011.
EBTS	IAFIS-DOC-01078-9.3: *Federal Bureau Of Investigation (FBI) Criminal Justice Information Services (CJIS) Electronic Biometric Transmission Specification (EBTS)*. https://www.fbibiospecs.org/docs/EBTS_v9_3_Final%2012_07_11_clean.pdf. Retrieved 2013-07-12.
JPEG	"T.81 : Information technology – Digital compression and coding of continuous-tone still images – Requirements and guidelines". http://www.itu.int/rec/T-REC-T.81. Retrieved 2011-01-12.
JPEG2K	"ISO/IEC 15444-1:2004 - Information technology -- JPEG 2000 image coding system: Core coding system". http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=27687. Retrieved 2009-11-01.
WSQ	"WSQ Gray-Scale Fingerprint Image Compression Specification" Version 3.1. https://www.fbibiospecs.org/docs/WSQ_Gray-scale_Specification_Version_3_1.pdf. Retrieved 2010-01-11.

Appendix A. **Dataset Makeup**

The study utilized fingerprint images based on the Basic Demonstration Model (BDM) fingerprints utilized in early tests of the FBI IAFIS system, and later used as the basis for the NIST SD-27 special database [SD27]. The ink card scan data and associated latent fingerprints were collected as a result of law enforcement activities and represents actual field data with collection dates ranging from 08/18/1973 through 04/12/1994. The original FD-249[6] fingerprint collection cards with these images were retrieved by NIST and rescanned at 1000 ppi by NIST personnel under controlled conditions. The images were scanned at 8 bits per pixel gray-scale using an FBI certified scanner and software (Appendix F complaint) and stored in a non-compressed format to ensure no compression anomalies are introduced into the original set.

As these latent impressions were taken from actual case work there was no effort to select images with any sort of demographic balance (pattern classifications, hand, etc.), such as what was done in [NISTIR 7778]. The selection of the samples used was based on the identity of a single latent fingerprint and not necessarily individual subjects. As such, 200 latent fingerprint impressions were selected from 175 subjects (i.e., some subjects contributed more than one finger). The latent fingerprints in SD-27 are classified into three quality categorizations (good, bad, ugly). These categorizations can be roughly defined as:

- Good: Well defined ridge structure, fingerprint area and well captured core.
- Bad: Some fragmentation of ridge structure, some discontinuity, but core mostly visible.
- Ugly: Highly fragmented ridge structure, poorly defined image area, core missing or undefined.

Also, due to some exemplars being selected for use in both matching and non-matching pairs, the selection and pairing process resulted in a dataset containing 200 latent impressions, 370 rolled impressions, and 371 flat impressions. In total, 30 rolled impressions and 29 flat impressions were used in both the IDENT and NON-IDENT pairings; however the actual pairings are unique and have not been duplicated. As such, there are 2,000 total unique pairings in the dataset used for this experiment.

In comparing a particular impression to that same impression, the same exact image/impression was used. For example, in the latent-to-latent match cases, one latent fingerprint image served as both the original uncompressed image as well as the compressed case image (see Case 1 in Figure 5). In the cases where one impression could not be used (i.e., comparing a latent impression to a rolled impression, such as Case 2 in Figure 5) different impressions from the same person were used in formulating the pair. For non-match data, an image from a similar pattern class was selected but from a different subject.

| Case 1 – Comparison of Same Impression | Case 2 – Comparison of Mixed Impressions |
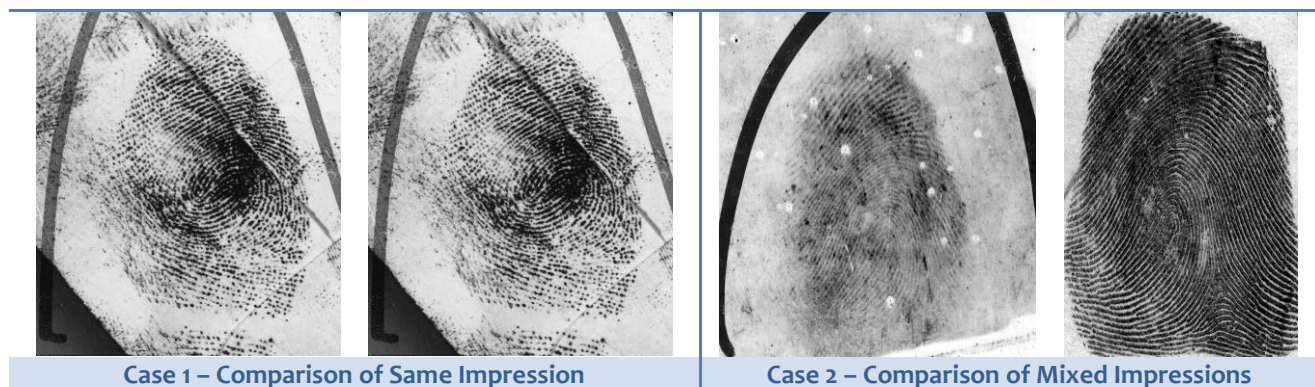
Figure 5 - Impression Comparison Examples

---

[6] The FD-249 card is a standard collection card typically used in the United States for controlled fingerprint collection in criminal casework.

Further information on the makeup of the data is provided in the tables in this section.

Table 21 - Ink Card Scan Data classification by Impression Type

Impression Type	Males	Females	Right	Left
Flat Single Finger	345	26	162	208
Rolled Single Finger	344	26	162	208
Latent Fingerprints	186	14	83	117

Table 22 - Gender Breakdown for Data

	Males	Females		White	Black	Hispanic	Asian	Native American
All Data	161	14		90	74	7	3	1

Table 23 - Age Breakdown for Data

	Under 18	18-24	25-29	30-34	35-39	40-44	45-49	50+	UNKNOWN
All Data	6	113	29	7	4	7	7	1	1

Table 24 - Other Metadata: Height and Weight

	5'0" – 5'5"	5'6" – 5'11"	6'0"+	UNKNOWN		<100 lbs	100 – 149 lbs	150 – 199 lbs	200 – 249 lbs	250+ lbs	UNKNOWN
All Data	15	113	46	1		0	48	103	19	4	1

Table 25 - Other Metadata: Eye Color

	Brown	Black	Blue	Green	Hazel	UNKNOWN
All Data	131	2	24	8	9	1

Table 26 - Image Geometry Data

Impression Type	Image Width (Pixels)				Image Height (Pixels)				Image Size (KB, Non-compressed)			
	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Rolled Single Finger (370 images)	1049	1041	549	1721	1192	1170	646	2012	1260.8	1254.9	389.5	2514.4
Flat Single Finger (371 images)	706	729	438	955	1285	1982	334	2140	938.2	693.8	198.1	1651.4
Latent (200 images)	1421	1442	828	1652	1379	1407	874	1557	1964.7	2017.2	1029.6	2533.9

Table 27 - Latent Quality Classification Data

	"Good"	"Bad"	"Ugly"
Latent Impressions (200 total)	61	67	72

Appendix B. **Equipment Used for Study**

1 x Commodity Router:
The router provides connectivity among the three workstations and network-attached storage (NAS) device, as well as providing remote access to the workstations and NAS for administration.

3 x Commodity Workstations:
The workstations are configured with 8 GB RAM, 300 GB HD, 64-bit operating system, FIXT software and data.

3 x 24 inch Monitors:
Three in-plane switching (IPS) panel monitors are connected via digital visual interface (DVI-D) and calibrated (see below) for optimal accuracy and consistency. The monitors were operated at their native resolution of 1920x1200, yielding a spatial geometry of approximately 94.3 ppi. Currently the state-of-art for stand-alone display technologies does not provide the capability for displaying fingerprints at anywhere near the native resolution of 500 ppi or 1000 ppi, and the best commercially available desktop displays operate at approximately 200 ppi. Because of this limitation zoom functionality in the test software was used to provide the ability to display the fingerprints at the closest possible resolution to the original image (10x), and the examiners were allowed to adjust the zoom up to approximately 50x.

1 x Network Attached Storage (NAS):
The NAS contains master copies of the FIXT software and data, as well as iterative copies of each stations logs/results (saved at the end of each session).

1 x Monitor Calibration Device:
The monitors were calibrated using a system which consists of a colorimeter paired with proprietary software designed specifically for use with the colorimeter and for the purpose of monitor calibration. The colorimeter is a sensor which provides an accurate measurement of colors as they actually appear on the monitor screen. During the calibration process, the colorimeter is physically attached to the monitor while the software displays a series of solid colors on the screen. The colorimeter measures the actual color values displayed on the monitor and then provides these measurements to the software. The software uses these measurements to calculate the difference between the color values as they are displayed on the monitor against the true color values within the software. The software then applies configuration changes to the system in order to correct the color values displayed by the monitor, ensuring accurate color reproduction. Due to the fact that each monitor, even of the same model, performs slightly differently in terms of the accuracy of color reproduction, this process was completed independently on each of the three aforementioned commodity workstations.