

NISTIR 7779

An Exploration of the Operational Ramifications of Lossless Compression of 1000 ppi Fingerprint Imagery

Shahram Orandi
John M. Libert
John D. Grantham
Kenneth Ko
Stephen S. Wood
Jin Chu Wu
Lindsay M. Petersen
Bruce Bandini

<http://dx.doi.org/10.6028/NIST.IR.7779>

NISTIR 7779

An Exploration of the Operational Ramifications of Lossless Compression of 1000 ppi Fingerprint Imagery

Shahram Orandi
John M. Libert
Kenneth Ko
Stephen S. Wood
Jin Chu Wu

*Information Access Division - Image Group
Information Technology Laboratories*

John D. Grantham
Systems Plus, Inc.

Lindsay M. Petersen
MITRE Corporation

Bruce Bandini
Booz Allen Hamilton, Inc.

<http://dx.doi.org/10.6028/NIST.IR.7779>

August 2012



U.S. Department of Commerce
Rebecca Blank, Acting Secretary

National Institute of Standards and Technology
Patrick D. Gallagher, Under Secretary of Commerce for Standards and Technology and Director

ACKNOWLEDGEMENTS

The authors wish to give special thanks to the following individuals and organizations for their support of this work:

Federal Bureau of Investigation for all their support throughout this study

T.J. Smith and the LA County Sheriff's Department

R. Michael McCabe, IDTP

Margaret Lepley, MITRE

In addition, we are grateful for and appreciate the guidance, support and coordination provided by Michael D. Garris, without whose help this study would not have been possible.

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.

EXECUTIVE SUMMARY

The criminal justice communities throughout the world exchange fingerprint imagery data primarily in 8-bit gray-scale and at 500 pixels per inch¹ (ppi) or equivalently 19.7 pixels per millimeter (ppmm). The Wavelet Scalar Quantization (WSQ) fingerprint image compression algorithm has been developed and maintained by the Federal Bureau of Investigation (FBI), Los Alamos National Laboratory and the National Institute for Standards and Technology (NIST) as the standard for compressing 500 ppi fingerprint imagery in the United States. WSQ is classified as a “lossy” compression algorithm. Lossy compression algorithms employ data encoding methods which discard (lose) some of the data in the encoding process in order to achieve an aggressive reduction in the size of the data being compressed. 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 users 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 1 for an example of image degradation from lossy compression).

The importance of latent images (images lifted from the crime scene, via uncontrolled collection) in criminal casework has been growing. Anecdotal evidence from latent fingerprint examiners has indicated that any fidelity loss as a result of compression should be avoided prior to review of those fingerprints by examiners and that even small increases in the amount of higher level detail may yield benefits. Because of this, most latent fingerprint images used in casework are either transferred in non-compressed form or compressed using a lossless algorithm where the resulting compressed representation of the image can be decompressed to yield an image exactly identical to the original.

While extensive effort has been put into standards and certification pathways for lossy fingerprint compression, lossless compression strategies have not been thoroughly examined.

This study examines various lossless compression algorithms and the relative advantages and disadvantages of each with respect to effective compression rates, compression throughput and decompression throughput. This study also examines different implementations of the same algorithm, as well as implementations of the same algorithm generated for different computer system architectures (e.g., 32-bit & 64-bit).

This study finds that wavelet-based compression algorithms such as JPEG 2000² generally yield better effective compression rates than the non-wavelet-based algorithms such as PNG (Portable Network Graphics). However, non-wavelet-based algorithms tend to have higher throughput (require less time to operate on data) than wavelet-based algorithms. This study also shows that while architectural differences do not yield much operational difference in terms of effective compression rates, such architectural differences do translate to significant differences in compression and decompression throughputs.

¹Resolution 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.

² The “2000” refers to the year of publication of the first edition of the image compression standard known as JPEG 2000. JPEG refers to Joint Photographic Experts Group.

VERSION HISTORY

Date	Activity

TABLE OF CONTENTS

1. Investigative Goals and Objectives	17
1.1. Background	17
1.2. Key Drivers and Mandates	18
2. Materials and Methods.....	19
2.1. Compression Algorithms	19
2.2. Compression Conditions	21
3. Analysis.....	22
3.1. Performance Metrics	22
3.2. Uncertainty Measurement	23
3.3. Algorithmic Comparisons	24
3.4. Architecture and Implementation Comparisons	28
4. Results.....	31
4.1. Investigative Goal 1: Examine Effective Compression Rates by Algorithm	31
4.2. Investigative Goal 2: Examine Compression Throughput by Algorithm	32
4.3. Investigative Goal 3: Examine Decompression Throughput by Algorithm.....	33
4.4. Investigative Goal 4: Examine Effective Compression Rate by Impression Type.....	34
4.5. Investigative Goal 5: Comparison of JPEG 2000 Implementation With Respect to Effective Compression Rate	35
4.6. Investigative Goal 6: Comparison of JPEG 2000 Implementation With Respect to Compression Throughput	36
4.7. Investigative Goal 7: Comparison of JPEG 2000 Implementation With Respect to Decompression Throughput.....	37
4.8. Investigative Goal 8: Comparison of Implementation Complexity	38
5. Conclusions	39
6. Discussions and Future Work	40
References	41
Publications and Reports	41
Standards.....	43
Appendix A. Dataset Makeup	45
Appendix B. Equipment Used for Study.....	50
Appendix C. Examination of Entropy	51
Procedure	51
Results	52

LIST OF TABLES

Table 1 - Abbreviations..... 15

Table 2 – Impression Types Selected for This Study 19

Table 3 - Compression Conditions21

Table 4 - Effective Compression Rates (Medians), Higher is Better..... 31

Table 5 – Effective Compression, Wilcoxon Signed Rank Test, alpha of 0.05 31

Table 6 - Compression Throughput (Median Time, seconds), Lower is Better 32

Table 7 - Compression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05 32

Table 8 - Decompression Throughput (Median Time, seconds)..... 33

Table 9 - Decompression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05..... 33

Table 10 - Algorithm Performance Rankings (1-Best through 6-Worst)..... 34

Table 11 - Effective Compression Rates (Medians) by Implementation, Higher is Better 35

Table 12 - Effective Compression by Implementation, Wilcoxon Signed Rank Test, at alpha = 0.05..... 35

Table 13 - Compression Throughput (Median Time, seconds) by Implementation, Lower is Better 36

Table 14 - Compression Throughput by Implementation, Wilcoxon Signed Rank Test, at alpha = 0.0033 36

Table 15 - Decompression Throughput (Median Time, seconds) by Implementation, Lower is Better 37

Table 16 - Decompression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05 37

Table 17 - Implementation Complexity 38

Table 18 - Ink Card Scan Data classification by Impression Type 45

Table 19 - Ink Card Scan Pattern Classification for Single Finger Images by Impression Type..... 45

Table 20 - Ink Card Scan Pattern Classification for Single Finger Images by Finger (Females)..... 46

Table 21 - Ink Card Scan Pattern Classification for Single Finger Images by Finger (Males) 46

Table 22 – Digital Live Scan Data Classification by Impression Type 46

Table 23 – Digital Live Scan Pattern Classification for Single Finger Images by Impression Type 47

Table 24 – Digital Live Scan Pattern Classification for Single Finger Images by Finger (Females) 47

Table 25 – Digital Live Scan Pattern Classification for Single Finger Images by Finger (Males) 47

Table 26 - Gender Breakdown for Data 48

Table 27 - Age Breakdown for Data 48

Table 28 - Other Metadata: Height and Weight 48

Table 29 - Other Metadata: Eye Color 48

Table 30 - Image Geometry Data..... 49

Table 31 – Mean measurements of entropy and fingerprint region of images 52

LIST OF FIGURES

Figure 1 - Example of Fidelity Degradation Due to Extreme Lossy Compression (JPEG 2000 at 800:1)..... 18

Figure 2 - Normal Fit Histogram, Effective Compression (PNG)..... 23

Figure 3 - Q-Q Normality Plot, Effective Compression (PNG) 23

Figure 4 - Effective Compression Rates (Medians) 25

Figure 5 - Compression Throughput (Median Time, seconds)..... 26

Figure 6 - Decompression Throughput (Median Time, seconds)..... 27

Figure 7 - Effective Compression Rate (Median Ratio, by Platform and Implementation)..... 28

Figure 8 - Compression Throughput (Median Time, seconds, by Platform and Implementation)..... 29

Figure 9 - Decompression Throughput (Median Time, seconds, by Platform and Implementation)..... 30

TERMS AND DEFINITIONS

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

Table 1 - Abbreviations

BMP	Bitmap File Format
bpp	Bits per pixel
CR	Compression Ratio
FBI	Federal Bureau of Investigation
IAFIS	Integrated Automated Fingerprint Identification System
IAI	International Association for Identification
ICER	Image Compression and Stereo Ranging
JPEG	Joint Photographic Experts Group – ISO/IEC committee developing standards for image compression – also used as the name of the CODEC developed in accordance with the standard specified by this body.
NBIS	NIST Biometric Image Software
NGI	Next Generation Identification
NIST	National Institute of Standards and Technology
OPJ	OpenJPEG’s JPEG 2000 CODEC
PNG	Portable Network Graphics
ppi	Pixels per inch
ppmm	Pixels per millimeter
PSNR	Peak Signal To Noise Ratio
RLE	Run Length Encoding
SIVV	Spectral Image Validation/Verification Metric
WSQ	Wavelet Scalar Quantization algorithm for compression of fingerprint imagery

ABSTRACT

This paper presents the findings of a study initially conducted to measure the operational impact of JPEG 2000 lossy compression on 1000 ppi fingerprint imagery at various levels of compression, but later expanded to include lossless compression. Lossless compression will have no impact on either Galton or non-Galton based features of a fingerprint since the compressed image is identical to the original once decompressed. The selection of a lossless compression algorithm can have operational implications in terms of effective compression rate and throughput; these implications are the focus. This study examines several such compression algorithms and compares them using criteria used to measure the effectiveness of the compression algorithm as well as its throughput using actual fingerprint imagery.

KEYWORDS

Fingerprint compression; 1000 ppi fingerprint imagery; JPEG 2000; JasPer; OpenJPEG; PNG; RLE; BMP; ICER; lossless compression

1. Investigative Goals and Objectives

In July of 2009 NIST in partnership with the FBI commenced an investigation on the use of JPEG 2000 [JPEG2K] for compressing fingerprint imagery. Part one of this investigation, described in [NISTIR7778], addressed JPEG 2000 when operating with a non-reversible/lossy filter. The study described in this paper complements [NISTIR 7778] by examining the performance of several compression algorithms including JPEG 2000 when operating in a lossless fashion with the following investigative goals:

1. **Examine Effective Compression Rates by Algorithm:** Assess the performance of the selected algorithms according to how much effective compression they yield for various impression types.
2. **Examine Compression Throughput by Algorithm:** Assess the performance of the selected algorithms with respect to the time needed to generate a compressed representation of the original image.
3. **Examine Decompression Throughput by Algorithm:** Assess the performance of the selected algorithms with respect to the time needed to reconstruct the original image from the compressed representation.
4. **Examine Effective Compression Rate by Impression Type:** Identify if a particular impression type impacts effective compression rates more than other types.
5. **Compare JPEG 2000 Implementations with Respect to Effective Compression Rate:** Identify any operational differences in effective compression rate between various implementations of the JPEG 2000 algorithm for the same given set of input images.
6. **Compare JPEG 2000 Implementations with Respect to Compression Throughput:** Identify any operational differences in compression throughput between various implementations of the JPEG 2000 algorithm for the same given set of input images.
7. **Compare JPEG 2000 Implementations with Respect to Decompression Throughput:** Identify any operational differences in decompression throughput between various implementations of the JPEG 2000 algorithm³ for the same given set of input images.
8. **Compare Implementation Complexity:** Examine and compare algorithm codebase size and complexity.

As an ancillary component of this study, the compression/decompression pathway was empirically verified to be truly lossless where the resulting compressed image stream would yield the exact original image, pixel-for-pixel.

1.1. Background

The criminal justice communities throughout the world exchange fingerprint imagery data primarily in 8-bit gray-scale and at 500 pixels per inch (ppi). The Wavelet Scalar Quantization (WSQ) [BRADLEY1], [BRADLEY2], [BRISLAWN], [HOPPER] image compression algorithm has been developed and maintained by the Federal Bureau of Investigation (FBI), Los Alamos National Laboratory (LANL) and the National Institute for Standards and Technology (NIST) as the standard for compressing 500 ppi fingerprint imagery in the United States. 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.

WSQ is a “lossy” compression algorithm. Lossy compression algorithms employ data encoding methods that discard (lose) some of the data in the encoding process in order to achieve an aggressive reduction in the size of the data being compressed. 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. Lossless compression algorithms, on the other hand, produce a compressed image that can be decompressed back to original form with no loss or change to the image. The disadvantage to lossless algorithms is that they produce compressed images that can be many times larger in file size than compressed images produced by lossy algorithms.

³ For purposes of the present study, both compression and decompression were performed by the same JPEG 2000 implementation. That is, there was no cross-over among implementations, e.g. encoding with one and decoding with another CODEC.

The WSQ algorithm allows users to specify how much compression to apply, where higher amounts of compression result in a more compact representation but greater loss in fidelity to the original image.

Figure 1 below shows an example of such image degradation and fidelity loss as a result of lossy compression. The WSQ Gray-Scale Fingerprint Image Compression Specification [WSQ] provides guidance for the acceptable amount of fidelity loss due to compression in order for the encoder and decoder to meet FBI certifications for 500 ppi fingerprint imagery. These certifications are designed to ensure adherence to the WSQ specification to ensure sufficient fidelity for admissibility of images in courts of law that have been processed by such encoders and decoders.



Figure 1 - Example of Fidelity Degradation Due to Extreme Lossy Compression (JPEG 2000 at 800:1)

A study conducted by the International Association for Identification (IAI) [FITZPATRICK] established 15:1 as a WSQ compression ratio that would retain acceptable image fidelity in 500 ppi fingerprint imagery. The study used the judgments of expert fingerprint examiners to measure the fidelity loss due to compression. In order to reduce bias due to subjectivity, multiple examiner decisions were used to build consensus. Utilizing examiners' opinions does not imply that automated fingerprint matcher performance is not an important criterion in a given biometric system, but it must be noted that 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.

1.2. Key Drivers and Mandates

In modernizing its environment as part of the Next Generation Identification (NGI) initiative, the FBI seeks to expand its ability to exchange fingerprints at 1000 ppi in an effort to improve upon the capacity of systems in fingerprint identification and verification tasks and meet the FBI mandate to:

- Protect the United States from terrorist attack, foreign intelligence operations and espionage
- Support federal, state, local and international partners in their efforts to prevent or reduce crime and violence
- Upgrade technology to support the FBI's missions

Toward meeting these goals, the FBI seeks to set guidance for the next generation encoders and decoders based on the JPEG 2000 compression standard [JP2K] in order to ensure interoperability, fidelity and admissibility for 1000 ppi images in courts of law in the criminal justice community. While this is the case for lossy compression, other CODECs might be considered for lossless compression of fingerprints for which the desire is to maintain absolute fidelity of the image to the original, yet achieve some savings of storage space.

In support of the FBI, the National Institute of Standards and Technology (NIST) conducted a study to determine an optimal compression approach that follows on the IAI study of WSQ compression for 500 ppi fingerprint imagery, to

build upon existing guidance for JPEG 2000 compression of fingerprint imagery, and to formulate a basis with which a normative compression guidance can be established in the ANSI/NIST standard for biometric data interchange [AN2011].

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 years, starting with automated fingerprint analysis which began in 1965.

The MITRE Corporation has developed an informative guidance that is widely recognized as the *de facto* standard guidance for utilizing JPEG 2000 for the compression of 1000 ppi fingerprint imagery in MTR-04B000022 [MTR]. While this document provides a compression profile for 1000 ppi fingerprint imagery using JPEG 2000, the guidance focuses on lossy compression of fingerprints. Unlike WSQ, JPEG 2000 supports both lossy functionality similar to WSQ (see 1.1), as well as a lossless mode. This study examines JPEG 2000 in lossless mode and compares its performance to other lossless compression algorithms.

2. Materials and Methods

A total of 1400 images were selected for this study from the NIST SD-27 special database [SD27]. These images included various impression types (see Table 2) and included multiple subjects (see Appendix A for more information on the makeup of the input data set). The images were then processed by the algorithms selected for this study. The algorithms selected were instrumented specifically for this study to enable the collection of timing data. Effective compression rate data was calculated from the resulting compressed files at a later time. No specific protocol was put in place for the ordering of images in the input set, however the images were processed in an arbitrarily stratified fashion in the order of case numbers 2, 4, 1, 3, 6, 5 and 7.

Each algorithm was tested with all the images independently and the images were processed sequentially without any delays or throttling of the processing stages.

Table 2 –Impression Types Selected for This Study

Case Number	Data Medium	Impression Type	Count
1	Ink Card Scan	Rolled single finger	200
2	Ink Card Scan	Flat single finger	200
3	Ink Card Scan	Slap-four finger	200
4	Ink Card Scan	Latent Lift Image	200
5	Digital Live Scan	Rolled single finger	200
6	Digital Live Scan	Flat single finger	200
7	Digital Live Scan	Slap-four finger	200
			Total: 1,400

2.1. Compression Algorithms

The focus of this study is to examine the effectiveness of various lossless compression algorithms on 1000 ppi fingerprint images. While there are many suitable algorithms, this study focuses on five: two implementations of JPEG 2000 (OpenJPEG and JasPer [JPEG2K]), PNG, RLE used in BMP, and ICER. JPEG 2000 (lossless) and PNG were included in this selection set because they are currently part of the existing standard for fingerprint data exchange [AN2011]. Although the ANSI/NIST standard also lists lossless JPEG, this format was not examined in this study due to anecdotal evidence indicating the low rate of acceptance of this compression format for 500 ppi fingerprint images. RLE is one of

the oldest and simplest compression algorithms and was included as a baseline due to its simple algorithm design and wide implementation availability. ICER is quite new, and was selected to represent a robust state-of-the art approach. All algorithms were tested on the same machine (See Appendix B for test machine configuration) to eliminate any computing platform bias.

2.1.1. JPEG 2000

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. In addition to providing a lossy approach, JPEG 2000 also provides for a lossless/reversible filter.

2.1.2. PNG

Portable Network Graphics (PNG) is an image compression standard and coding system that was created by the PNG Working Group and later accepted as a standard in 2004 (ISO/IEC 15948:2004) [ISO/IEC]. PNG was primarily created to improve upon and replace GIF (Graphics Interchange Format) as an open image file format not requiring a patent license. PNG utilizes a 2-stage compression process where a prediction filter is applied to the image in an attempt to make the data more compressible in the ultimate compression stage. PNG is a very flexible compression algorithm supporting palette-based color images, grayscale images, and RGB images.

2.1.3. RLE Used in BMP

Run-Length Encoding (RLE) is a very simple form of data compression used by the BMP file format in which sequences of repetitive data are stored as a single data value with an associated repeat count for that value, rather than as the original sequence. This compression algorithm is most useful on data that contains many such repeating sequences of the same value, for example a largely blank image containing only white pixels. RLE is not effective with files that exhibit any significant entropy⁴ in structure, and frequently the file size of such data can greatly increase after being compressed with RLE. This is due to the addition of count values to the byte stream when few data are actually repeated, resulting in more data added than removed or compressed.

2.1.4. ICER

ICER (Image Compression and Stereo Ranging) [KIELY1],[KIELY2] is a modern wavelet-based image data compression algorithm specially designed to meet the needs of deep-space exploration applications such as the Mars Exploration Rover (MER) where state-of-the-art data compression effectiveness and efficiency are primary concerns rather than catering to general purpose applications. ICER provides lossless and lossy compression modes. For the purposes of this study it was examined in lossless mode. While ICER is not generally available for commercial use, it has been included as a reference modern high-end compression algorithm, just as RLE has been included in this study to provide a reference low-end compression algorithm.

⁴ Appendix C provides a comprehensive description of entropy as interpreted in this study.

2.2. Compression Conditions

Per the investigative goals described in Section 1, this study’s experimental goals can be differentiated into two overarching tracks each consisting of 6 compression conditions (see Table 3 below). The first track focuses on the comparison of different compression algorithms (inter-algorithm conditions) while the second track focuses on the different implementations of the same algorithm (intra-algorithm, JPEG-2000 specifically).

Table 3 - Compression Conditions

Experimental Condition (Inter-Algorithm)	Experimental Conditions (Intra-Algorithm)
Non-Compressed	OpenJPEG v.1.3, 32-bit
RLE (32-bit)	OpenJPEG v.1.3, 64-bit
PNG (32-bit)	OpenJPEG v.1.4, 32-bit
JPEG 2000 (OpenJPEG v.1.4, 32-bit)	OpenJPEG v.1.4, 64-bit
JPEG 2000 (JasPer 1.900.1, 32-bit)	JasPer 1.900.1 32-bit
ICER (32-bit)	JasPer 1.900.1 64-bit

3. Analysis

3.1. Performance Metrics

In the case of compression performance measurement, there are three key factors in establishing overall performance. These factors include the effective compression ratio of the algorithm, the effective throughput of the algorithm when it compresses the image, and the effective throughput of the algorithm when it decompresses the image.

3.1.1. Effective Compression Ratio Measurement

The effective compression ratio is measured by how much size reduction is obtained relative to the original non-compressed image by applying the algorithm being studied. In examining effective compression across multiple test images of various sizes and impression type, as was the case in this study, simply comparing resulting image sizes is not practical as the input images are not of homogenous geometry prior to compression (See Table 30). A simple and common metric of effective compression is defined by calculating the ratio of the non-compressed (raw) image size to the compressed image size as defined below:

$$\text{Effective Compression} = \frac{\text{Raw File Size (Bytes)}}{\text{Compressed File Size (Bytes)}}$$

3.1.2. Effective Compression Throughput

As with other metrics, the effective compression throughput of an algorithm can be measured in many ways such as including or excluding IO overhead and disk cache effectiveness. For the scope of this study a simple clock-based approach on a dedicated machine was employed. This is described further in Appendix B. The compression time was measured by instrumenting each of the compression algorithms to record the difference in the time interval from launch to when the compressed image has been saved onto a hard disk. No attempt was made to exclude and/or measure disk IO overhead or to measure any impact of built in disk caching mechanism on compression throughput. All testing was performed on a single machine to avoid any bias introduced by differences in architecture and configuration between multiple machines.

3.1.3. Effective Decompression Throughput

The effective decompression throughput of an algorithm was measured similarly to the measurement of effective compression throughput. The decompression time was measured by instrumenting each of the decompression algorithms to record the difference in the computer's real-time-clock from launch to when the image had been successfully decompressed and saved onto a hard disk. No attempt was made to exclude and/or measure disk IO overhead or to measure any impact the built in disk caching mechanism has on decompression throughput. All testing was performed on a single machine to avoid any bias introduced by differences in architecture and configuration between multiple machines.

3.2. Uncertainty Measurement

3.2.1. Normality Test

Examination of the distribution of data in the normal probability plot [CHAMBERS] and the Q-Q Plot on a selected subset of data from this study (effective compression ratio performance of the PNG algorithm) in Figure 2 and Figure 3 shows that it is not possible to determine the distribution of the data and therefore parametric analysis methods are not suitable. Because of this we accept the median as the comparison statistic and utilize a bootstrap method which does not rely on the distribution of the underlying data to assess the uncertainty of the medians estimated from the observed sample measurements.

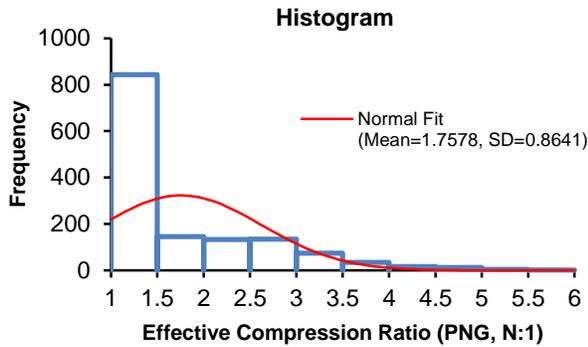


Figure 2 - Normal Fit Histogram, Effective Compression (PNG)

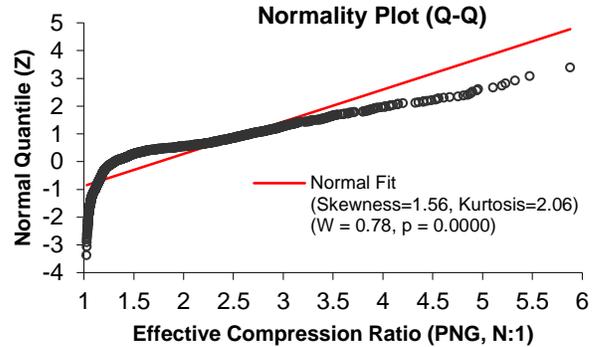


Figure 3 - Q-Q Normality Plot, Effective Compression (PNG)

3.2.2. Bootstrap Procedure

Given the observed non-normality of the measurements, we employ the sample median of $N=1400$ measurements of each compression metric (or $N=200$ measurements for each of seven different impression types) as our comparison statistic. We estimate the uncertainty, i.e., standard error, of the median using a bootstrap procedure [WU1], [WU2], [WU3] wherein we re-compute the median for each of 1,000 samples of size N , sampled randomly with replacement from the original N measurements. The 95 % confidence interval for the median statistic is then taken as the 0.025th and 0.975th quantiles of the distribution of median replicates. The sample medians and their upper and lower confidence limits are shown in figures 4 – 9 in sections below.

3.2.3. Hypothesis Testing

Hypothesis testing is performed using a non-parametric method, specifically the Wilcoxon Signed Rank Test [WILCOXON], [HOLLANDER]. Differences among the various compression processes are tested for each of three metrics. We test for differences in each of three metrics between every pair of compression processes without regard to directionality of differences (two-tailed test). We have two sets of compression conditions (Inter-Algorithm and Intra-Algorithm described in Section 2.2) under test yielding for each set 15 pairwise comparisons for each of the three comparison metrics (effective compression rate, compression throughput and decompression throughput). We also perform the pairwise comparison tests for each of the $p = 10$ fingerprint groupings including All Data, Latent, Ink Card Scan All, Ink Card Scan Rolled, Ink Card Scan Flat, Ink Card Scan Slap, Digital Live Scan All, Digital Live Scan Rolled, Digital Live Scan Flat, and Digital Live Scan Slap. That is, for each pair of compression treatments, we wish to test the null hypothesis, H_0 , that the median difference of pairwise measurements is zero. We reject the null hypothesis if we find the probability of its truth to be less than the Type I error rate, i.e. the probability of incorrectly rejecting the null hypothesis. Typically, the acceptable Type I error rate is set at 5%, i.e. $\alpha = 0.05$. If we are able to reject the null hypothesis, we accept the alternative hypothesis, H_1 , that the median of N pairwise differences in measurements is not equal to zero. Thus, for the present experiments, the null hypotheses, H_0 , and alternative hypotheses, H_1 , for the $n=15$ pairwise comparisons may be stated mathematically as:

$$\left\{ H_0^{c,k,g} : \text{Med}(m_{i,c,k,g} - m_{i,d,k,g}) = 0 \Big|_{c,k,g}^{i=1\dots N; c=1\dots 6; d=1\dots 6(d \neq c); k=1\dots 3; g=1\dots 9} \right\}$$

$$\left\{ H_1^{c,k,g} : \text{Med}(m_{i,c,k,g} - m_{i,d,k,g}) \neq 0 \Big|_{c,k,g}^{i=1\dots N; c=1\dots 6; d=1\dots 6(d \neq c); k=1\dots 3; g=1\dots 9} \right\}$$

where m in the expression denotes a metric further specified by subscripts; i designates a pair of measurements 1... N ; c designates the compression method 1...6; k designates metric 1... 3; and g designates fingerprint groupings 1... 9. Each of the paired tests is independent of other tests and the procedure is performed separately for each metric and for each group of fingerprint images as specified above. Moreover, in a similar analysis, we compare various implementations of the JPEG 2000 CODEC.

The Wilcoxon Signed Rank test examines differences between pairwise measurements without requiring the assumption of normality. This test is analogous to the pairwise t-test used to compare pairs of measurements having distributions known to satisfy the assumptions of normality. Even as subsequent tables and graphs present sample medians, the hypothesis test comparisons are made between pairs of measurements and not between distributions.

3.3. Algorithmic Comparisons

The first phase of this study entailed comparison of the various algorithms selected for this study and their relative differences. These differences can be summarized as the differences in the effectiveness of the algorithm in terms of compressing the image relative to the raw file size, the amount of time needed by the algorithm to compress the image and the amount of time needed by the algorithm to decompress that image.

3.3.1. Effective Compression Rate Comparison

Comparison of median values for effective compression ratios shows that the wavelet-based compression algorithms (ICER and JPEG 2000) hold a definitive advantage over non-wavelet-based compression algorithms (PNG, RLE) in compressing mixed image types as shown in Figure 4. The older RLE algorithm in the BMP format appears to be negatively impacted by the high entropy of fingerprint data where it incurs a negative compression penalty; that is, the image grows in size after compression relative to the original non-compressed file. One exception emerges however, in that Digital Live Scan Rolled images achieve higher compression ratios using PNG in contrast to wavelet-based algorithms. Examination of image entropies shows that Digital Live Scan images typically contain less entropy relative to Ink Card Scan images, particularly in the case of untextured backgrounds. This renders these images more compressible using traditional (non-wavelet-based) algorithms.

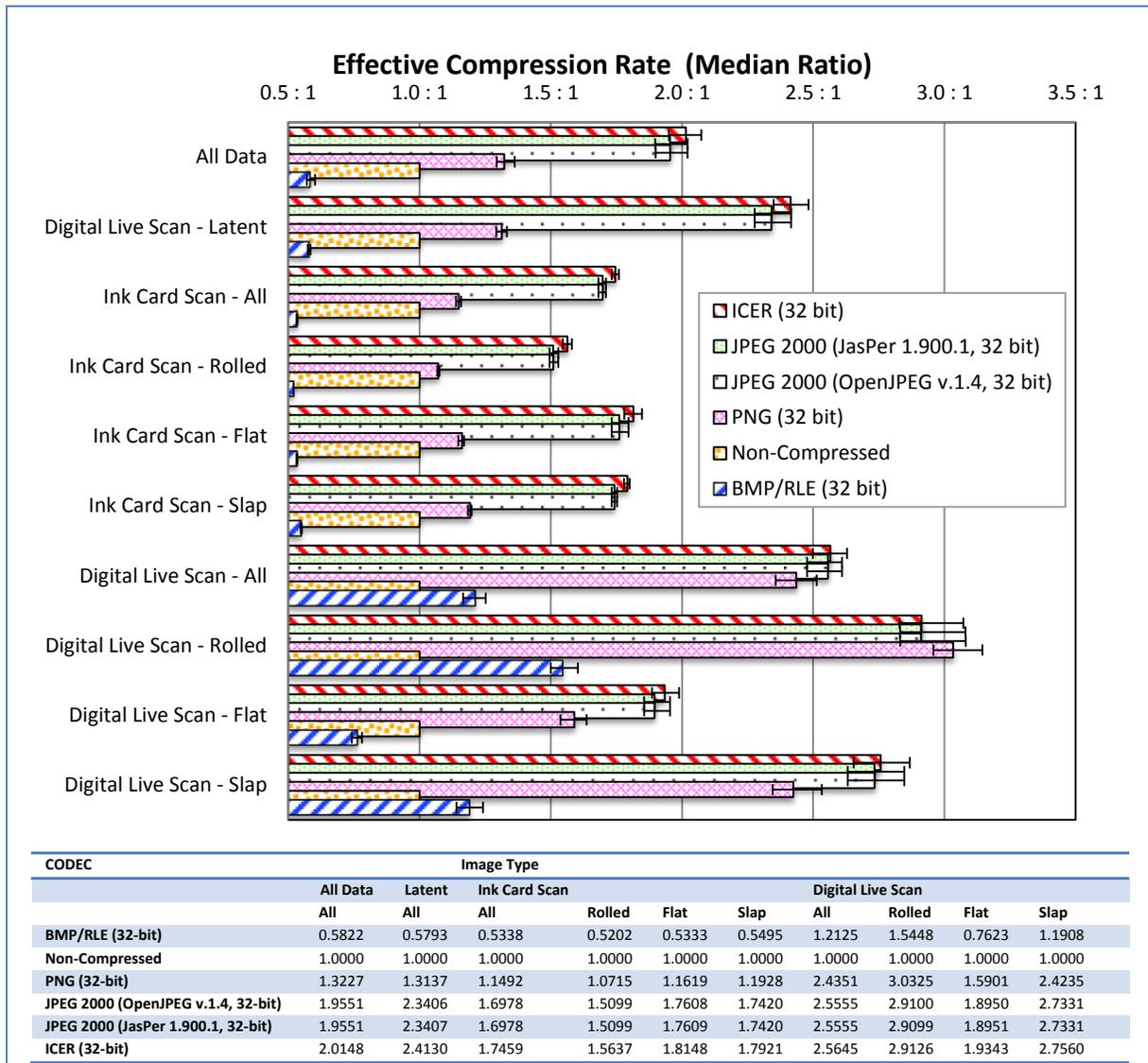


Figure 4 - Effective Compression Rates (Medians)

3.3.2. Compression Throughput Comparison

In addition to effective compression ratio, the time an algorithm takes to create a compressed representation of an image can also be a significant factor in the selection of an algorithm. The most notable example here is mobile applications where image compression may take place on a device with modest computational resources. It should also be noted that in common biometric processing workflows, the compression process typically occurs once after image acquisition, while the image may be decompressed one or more times; therefore the load incurred during compression is a transient one. Comparison of median values for effective compression time has been provided in Figure 5. Throughput data shows that while the compression time varies heavily by algorithm implementation, there is some stratification both by algorithm and image geometry (size). In the case of JPEG 2000, the two implementations of JPEG 2000 (JasPer and OpenJPEG) show significant differences in algorithm compression throughput. Examination of algorithm throughputs does not show a consistent advantage towards either wavelet or non-wavelet-based compression algorithms, but does show that compression throughput can vary greatly by specific implementations of the algorithm.

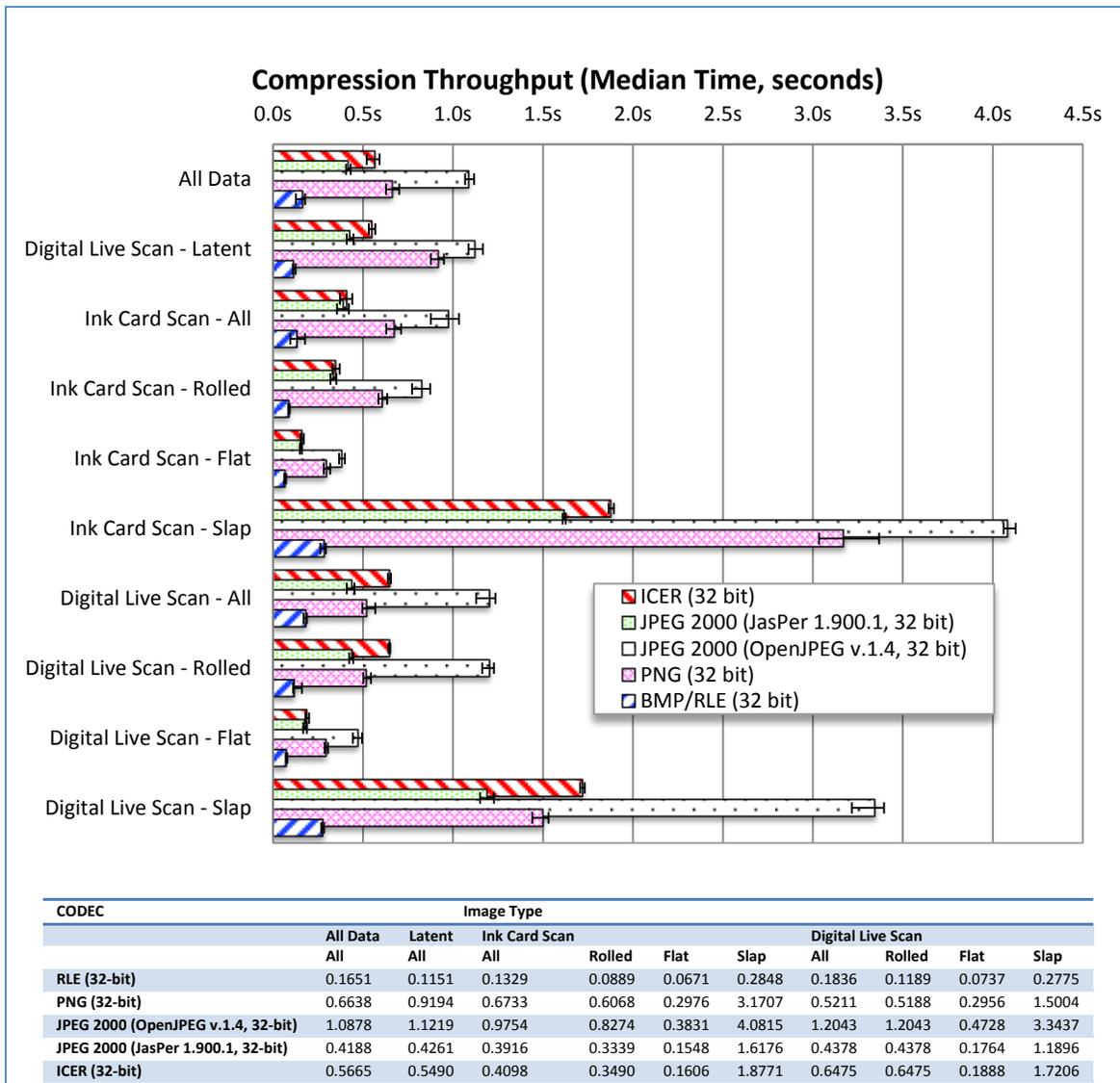


Figure 5 - Compression Throughput (Median Time, seconds)

3.3.3. Decompression Throughput Comparison

Another important factor in algorithm performance is the amount of time the algorithm takes to decompress the image for processing and/or viewing. The decompression time can be a significant factor in the selection of an algorithm as it can occur multiple times in the lifecycle of the biometric record. Comparison of median values for effective decompression time has been provided in Figure 6. Throughput data shows that while the compression time varies heavily by algorithm implementation, the non-wavelet-based compression algorithms (PNG, RLE) show an advantage in decompression throughput. This advantage is most dramatic with PNG. In the case of JPEG 2000, the two implementations of JPEG 2000 (JasPer and OpenJPEG) show significant differences in algorithm decompression throughput suggesting differing levels of optimization between specific implementations of the algorithm.

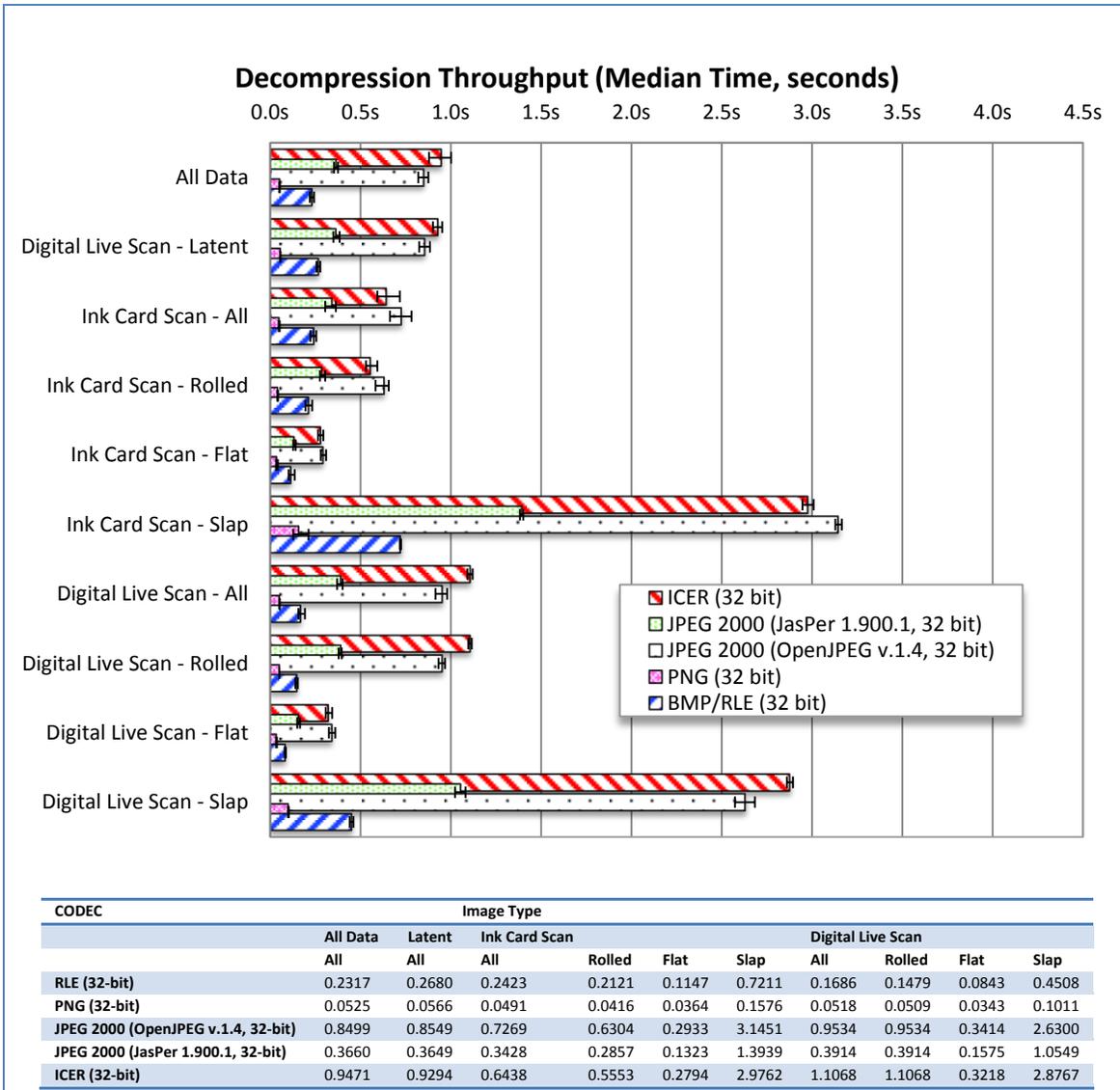


Figure 6 - Decompression Throughput (Median Time, seconds)

3.4. Architecture and Implementation Comparisons

The second phase of this study focused on the JPEG 2000 algorithm and examined the relative differences between the various implementations and architectural configurations of this algorithm. Again, as with section [3.3] the differences are summarized as the differences in the effectiveness of the algorithm in terms of compressing the image relative to the raw (non-compressed) file size, the amount of time needed by the algorithm to compress the image and the amount of time needed by the algorithm to decompress that image.

3.4.1. Effective Compression Rate Comparison

Comparison of median values for effective compression ratios shows that the various implementations of JPEG 2000 yield nearly identical compression rates.

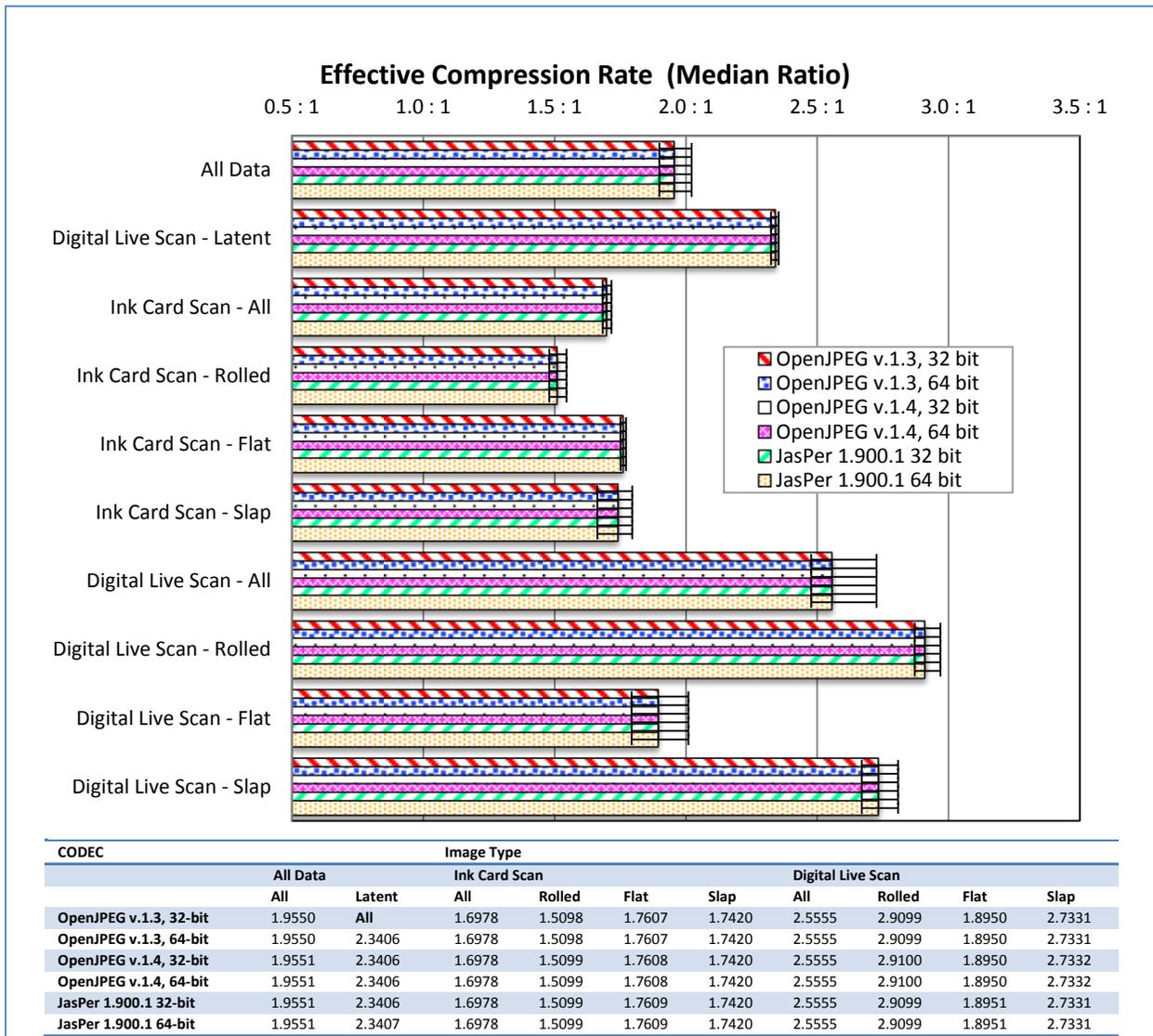


Figure 7 - Effective Compression Rate (Median Ratio, by Platform and Implementation)

3.4.2. Compression Throughput Comparison

Comparison of median values for effective compression time as shown in Figure 8 shows that platform optimization and evolution can lead to significant improvements in speed as evident from comparing OpenJPEG v.1.3 and v.1.4. Architectural differences however are not evident. For example, the JasPer implementation shows a performance penalty when compiled for a 64-bit architecture, while OpenJPEG appears to have performance gains when operating in a 64-bit environment.

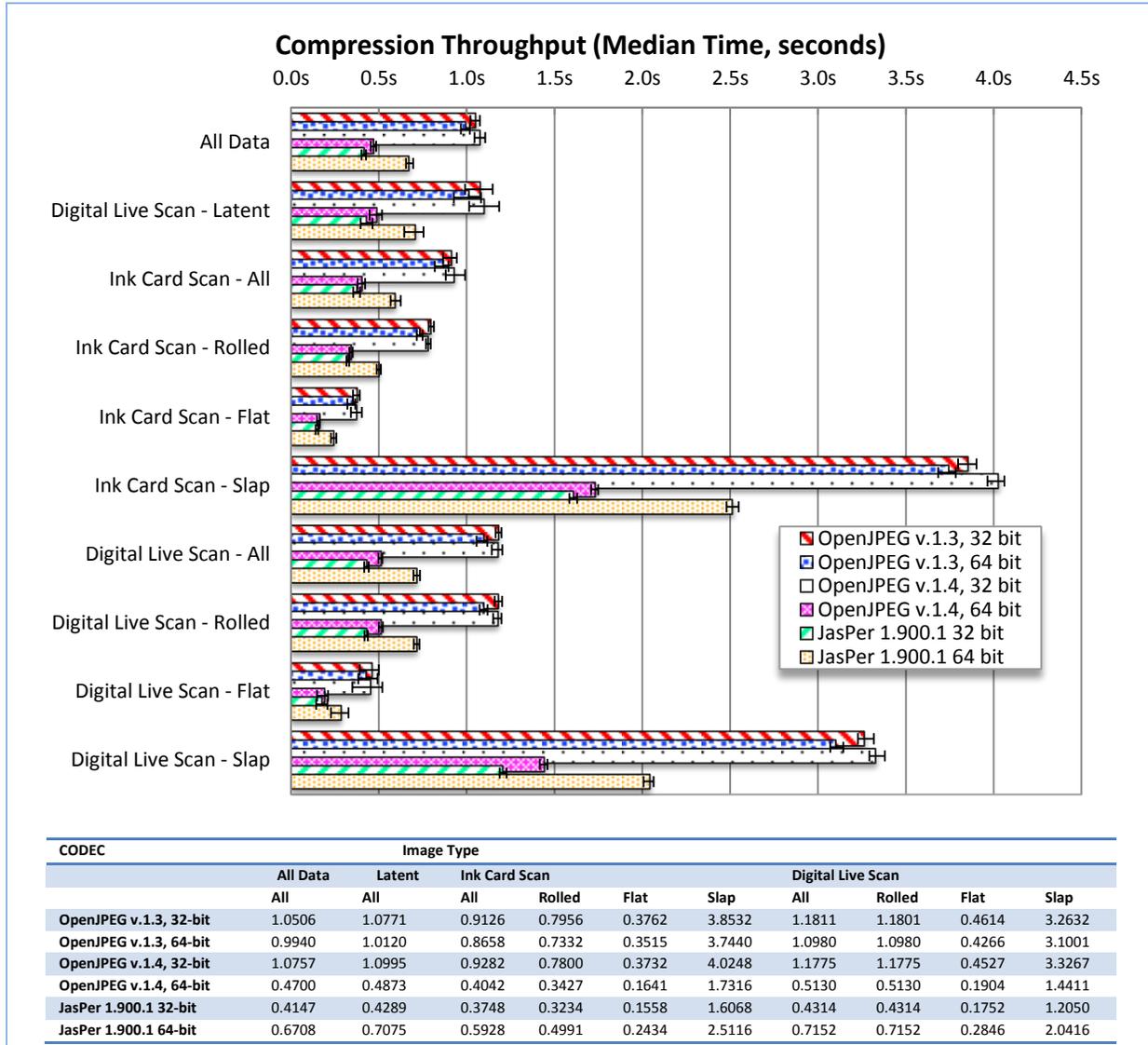


Figure 8 - Compression Throughput (Median Time, seconds, by Platform and Implementation)

3.4.3. Decompression Throughput Comparison

Comparison of median values for effective decompression time as shown in Figure 9 shows a similar pattern to the data collected in compression throughput measurements from Figure 8. Again, platform optimization and evolution appear to lead to significant improvements in decompression speed as evident from comparing OpenJPEG v.1.3 and v.1.4. Architectural differences do not appear to significantly or consistently impact decompression times. For example, the JasPer implementation shows a throughput penalty when operating in 64-bit compilation, while OpenJPEG appears to have performance gains when operating in 64-bit compilation.

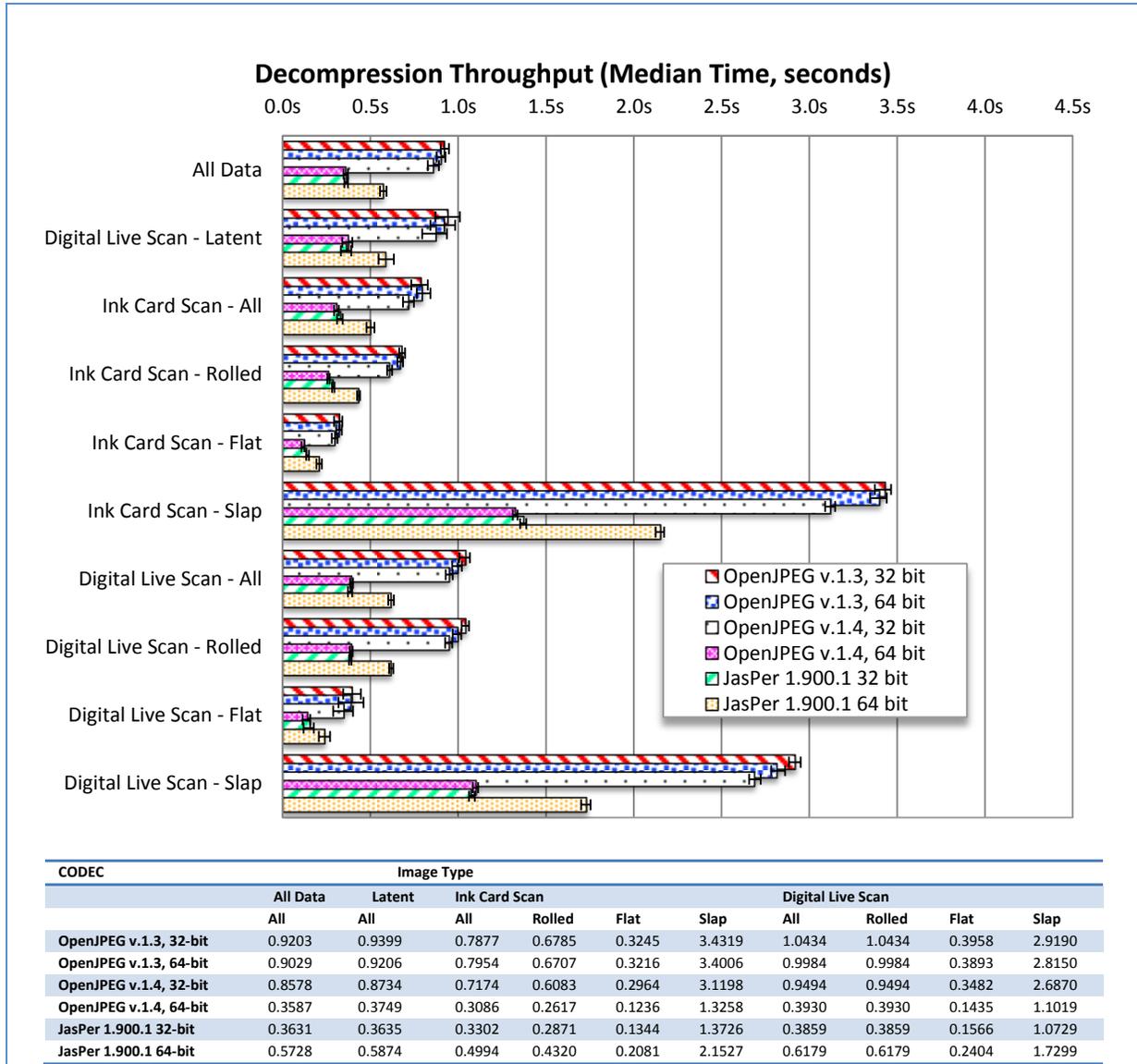


Figure 9 - Decompression Throughput (Median Time, seconds, by Platform and Implementation)

4. Results

4.1. Investigative Goal 1: Examine Effective Compression Rates by Algorithm

One of the primary drivers in the selection of a compression ratio is the effective compression rate obtained using that algorithm on the data of interest. The first investigative goal of this study is to determine if any particular algorithm holds a definitive advantage versus the others, and if the measured advantage is statistically significant.

4.1.1. Investigative Analysis 1

Based on the median data (Table 4), ICER provides the best compression rate for mixed image types (“All Data”) with a median ratio of 2.0148:1 followed by the implementations of JPEG 2000. Pair-wise examination of the data using Wilcoxon Signed Rank Test with a significance level (alpha) of 0.05 shows that the differences between ICER and the two JPEG 2000 algorithms are closer than other probabilities to the Type I error threshold of 0.05 for the Digital Live Scan case, but are still significantly different (Table 5). Hence the null hypothesis is rejected for this comparison as well as for all other comparisons for which the probability of incorrectly rejecting the null hypothesis are well below the 0.05 level. The data also shows that ICER and JPEG 2000 yield the best compression performance for every case except for Digital Live Scan Rolled imagery where PNG outperforms all other algorithms.

Table 4 - Effective Compression Rates (Medians), Higher is Better

CODEC	Image Type									
	All Data	Latent	Ink Card Scan				Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
RLE (32-bit)	0.5822	0.5793	0.5338	0.5202	0.5333	0.5495	1.2125	1.5448	0.7623	1.1908
Non-Compressed	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
PNG (32-bit)	1.3227	1.3137	1.1492	1.0715	1.1619	1.1928	2.4351	3.0325	1.5901	2.4235
JPEG 2000 (OpenJPEG v.1.4, 32-bit)	1.9551	2.3406	1.6978	1.5099	1.7608	1.7420	2.5555	2.9100	1.8950	2.7331
JPEG 2000 (JasPer 1.900.1, 32-bit)	1.9551	2.3407	1.6978	1.5099	1.7609	1.7420	2.5555	2.9099	1.8951	2.7331
ICER (32-bit)	2.0148	2.4130	1.7459	1.5637	1.8148	1.7921	2.5645	2.9126	1.9343	2.7560

Table 5 – Effective Compression, Wilcoxon Signed Rank Test, alpha of 0.05

Case	Comparison Pair		Image Type									
	CODEC 1	CODEC 2	All Data	Latent	Ink Card Scan				Digital Live Scan			
			All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
1	Non-Compressed	RLE	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	Non-Compressed	PNG	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	Non-Compressed	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	Non-Compressed	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	Non-Compressed	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	RLE	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	RLE	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	PNG	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	PNG	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	PNG	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	J2K (OpenJPEG v.1.4)	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0013	<0.0001	<0.0001	<0.0001
14	J2K (OpenJPEG v.1.4)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0400	<0.0001	<0.0001
15	J2K (JasPer 1.900.1)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0426	<0.0001	<0.0001

4.1.2. Investigative Result 1

Based on the data, the wavelet algorithms provide consistently better performance than their non-wavelet-based counterparts with results that are statistically significant. ICER leads the pack, followed by JPEG 2000. The only exception to this was the Live Scan Rolled image case. It is hypothesized that the lower entropy of this particular image type may make these images more effectively compressible by PNG. RLE on the other hand seems to suffer a significant penalty with high entropy images such as Ink Card Scan where the resulting compressed representation is larger than the original non-compressed images.

4.2. Investigative Goal 2: Examine Compression Throughput by Algorithm

Another important driver in the selection of a compression ratio is throughput of the algorithm in terms of the time it takes to generate the compressed representation of data of interest. Compression typically only occurs once upon the collection and packaging of the biometric sample and therefore it is not perceived as a source of any significant recurring load on the processing pathways of the typical biometric system. Compression time however can be a significant decision driver for devices with modest computational ability such as hand-held or embedded devices where even a transient load may have very significant impact on the workflow and usability of such devices. The second investigative goal of this study is to determine if any particular algorithm holds a definitive advantage in terms of compression throughput, and if the measured advantage is statistically significant.

4.2.1. Investigative Analysis 2

Based on the median data (Table 6), RLE provides the best compression throughput⁵ for mixed image types (“All Data”) with a median ratio of 0.1651 seconds, followed by the JasPer implementation of JPEG 2000 and PNG. Pair-wise examination of the data using Wilcoxon Signed Rank Test with a significance level (alpha) of 0.05 shows that all comparisons were statistically significant for all comparison pairs and image types, i.e. all probabilities are below the 0.05 level.

Table 6 - Compression Throughput (Median Time, seconds), Lower is Better

CODEC	Image Type									
	All Data	Latent	Ink Card Scan				Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
RLE (32-bit)	0.1651	0.1151	0.1329	0.0889	0.0671	0.2848	0.1836	0.1189	0.0737	0.2775
PNG (32-bit)	0.6638	0.9194	0.6733	0.6068	0.2976	3.1707	0.5211	0.5188	0.2956	1.5004
JPEG 2000 (OpenJPEG v.1.4, 32-bit)	1.0878	1.1219	0.9754	0.8274	0.3831	4.0815	1.2043	1.2043	0.4728	3.3437
JPEG 2000 (JasPer 1.900.1, 32-bit)	0.4188	0.4261	0.3916	0.3339	0.1548	1.6176	0.4378	0.4378	0.1764	1.1896
ICER (32-bit)	0.5665	0.5490	0.4098	0.3490	0.1606	1.8771	0.6475	0.6475	0.1888	1.7206

Table 7 - Compression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05

Case	Comparison Pair		Image Type										
			All Data	Latent	Ink Card Scan				Digital Live Scan				
	CODEC 1	CODEC 2	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap	
1	Non-Compressed	RLE	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	Non-Compressed	PNG	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	Non-Compressed	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	Non-Compressed	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	Non-Compressed	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	RLE	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	RLE	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	PNG	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	PNG	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	PNG	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	J2K (OpenJPEG v.1.4)	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
14	J2K (OpenJPEG v.1.4)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
15	J2K (JasPer 1.900.1)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

4.2.2. Investigative Result 2

The RLE (using BMP) algorithm provides it with the greatest throughput advantage for the mixed image type case (“All Data”). The JasPer implementation of JPEG 2000 and ICER follow RLE in throughput performance. It is hypothesized that the simplicity of the RLE’s compression is most likely responsible for the very high throughput, albeit at the cost of far less effective compression yield (See 4.1.1). The ordering of throughput is consistent with all image type cases, and statistically significant for every case.

⁵ It should be noted that the Non-Compressed case was omitted from this ranking. The Non-Compressed case, if taken in literal context with the other algorithms, would rank as best in compression throughput due to an elapsed compression time of 0.

4.3. Investigative Goal 3: Examine Decompression Throughput by Algorithm

Other than compression ratio, the decompression throughput of an algorithm is the next most important factor in terms of the time it takes to reconstruct the original image from its compressed representation (decompression). Image decompression is expected to occur at least one time along the processing pathway of a biometric system, and possibly many times throughout the lifecycle of that biometric sample; therefore it may be a significant decision driver. This non-transient load may again be a significant factor with portable devices. The third investigative goal of this study is to determine if any particular algorithm holds a definitive advantage versus the others in terms of decompression throughput, and if the measured advantage is statistically significant.

4.3.1. Investigative Analysis 3

Based on the median data (Table 8), PNG provides the best decompression throughput for mixed image types case (“All Data”) with a median of 0.0525 seconds, followed by RLE and the JasPer implementation of JPEG 2000. Pair-wise examination of the data using Wilcoxon Signed Rank Test with an alpha of 0.05 shows that all comparisons were statistically significant for all comparison pairs and image types, with probabilities of a Type I error well below the 0.05 level.

Table 8 - Decompression Throughput (Median Time, seconds)

CODEC	Image Type									
	All Data	Latent	Ink Card Scan				Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
RLE (32-bit)	0.2317	0.2680	0.2423	0.2121	0.1147	0.7211	0.1686	0.1479	0.0843	0.4508
PNG (32-bit)	0.0525	0.0566	0.0491	0.0416	0.0364	0.1576	0.0518	0.0509	0.0343	0.1011
JPEG 2000 (OpenJPEG v.1.4, 32-bit)	0.8499	0.8549	0.7269	0.6304	0.2933	3.1451	0.9534	0.9534	0.3414	2.6300
JPEG 2000 (JasPer 1.900.1, 32-bit)	0.3660	0.3649	0.3428	0.2857	0.1323	1.3939	0.3914	0.3914	0.1575	1.0549
ICER (32-bit)	0.9471	0.9294	0.6438	0.5553	0.2794	2.9762	1.1068	1.1068	0.3218	2.8767

Table 9 - Decompression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05

Case	Comparison Pair		Image Type									
	CODEC 1	CODEC 2	All Data	Latent	Ink Card Scan				Digital Live Scan			
			All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
1	Non-Compressed	RLE	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	Non-Compressed	PNG	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	Non-Compressed	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	Non-Compressed	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	Non-Compressed	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	RLE	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	RLE	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	RLE	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	PNG	J2K (OpenJPEG v.1.4)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	PNG	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	PNG	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	J2K (OpenJPEG v.1.4)	J2K (JasPer 1.900.1)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
14	J2K (OpenJPEG v.1.4)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
15	J2K (JasPer 1.900.1)	ICER	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

4.3.2. Investigative Result 3

PNG’s advantage in decompression throughput is dramatic and at times an order of magnitude better than the next closest performer. While PNG’s compression effectiveness may lag behind its wavelet-based counterparts (JPEG 2000 and ICER), the performance advantages in decompression may make it a suitable algorithm for asymmetric use cases where decompression performance is critical.

4.4. Investigative Goal 4: Examine Effective Compression Rate by Impression Type

Another factor in the selection of a compression algorithm is applicability of that algorithm to the taxonomy of the data that is expected as part of the workflow and processes. If the data has skewed taxonomy (i.e., only Flat images taken by a Digital Live Scan device) it may be beneficial to select a compression algorithm that performs better given that data. The fourth investigative goal of this study is to determine if any particular algorithm holds a definitive advantage versus the others in terms of performance by the type of image processed.

4.4.1. Investigative Analysis 4

The algorithms in this study were ranked in terms of performance for the various criteria of interest (effective compression ratio, compression throughput and decompression throughput). The results are provided in Table 10 below. Examination of rankings across the columns of Table 10 shows that rankings/stratification of algorithms appear to be relatively stable and consistent by various image types. There are cases where positions of algorithms with similar performance rankings switch, but this change in ranking is usually limited to algorithms with adjacent performance rankings. The notable exception to this is PNG's effective compression ratio performance with Digital Live Scan Rolled images where the algorithm jumps from a typical ranking of 4th place to 1st place relative to the other algorithms. For reference purposes, the non-compressed case has been included in the throughput rankings. The non-compressed cases here represent the system baseline of zero time required.

Table 10 - Algorithm Performance Rankings (1-Best through 6-Worst)

Factor	CODEC	Image Type											
		All Data		Latent		Ink Card Scan				Digital Live Scan			
		All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap		
Effective Ratio	Non-Compressed	5	5	5	5	5	5	6	6	5	6		
	BMP/RLE (32 bit)	6	6	6	6	6	6	5	5	6	5		
	PNG (32 bit)	4	4	4	4	4	4	4	1	4	4		
	JPEG 2000 (OpenJPEG v.1.4, 32 bit)	3	3	3	3	3	3	2	3	3	2		
	JPEG 2000 (JasPer 1.900.1, 32 bit)	2	2	2	2	2	2	3	4	2	3		
	ICER (32 bit)	1	1	1	1	1	1	1	2	1	1		
	Compression Throughput	Non-Compressed	1	1	1	1	1	1	1	1	1	1	
BMP/RLE (32 bit)	2	2	2	2	2	2	2	2	2	2			
PNG (32 bit)	5	5	5	5	5	5	4	4	5	4			
JPEG 2000 (OpenJPEG v.1.4, 32 bit)	6	6	6	6	6	6	6	6	6	6			
JPEG 2000 (JasPer 1.900.1, 32 bit)	3	3	3	3	3	3	3	3	3	3			
ICER (32 bit)	4	4	4	4	4	4	5	5	4	5			
Decompression Throughput	Non-Compressed	1	1	1	1	1	1	1	1	1	1		
	BMP/RLE (32 bit)	3	3	3	3	3	3	3	3	3	3		
	PNG (32 bit)	2	2	2	2	2	2	2	2	2	2		
	JPEG 2000 (OpenJPEG v.1.4, 32 bit)	5	5	6	6	6	6	5	5	6	5		
	JPEG 2000 (JasPer 1.900.1, 32 bit)	4	4	4	4	4	4	4	4	4	4		
	ICER (32 bit)	6	6	5	5	5	5	6	6	5	6		

4.4.2. Investigative Result 4

Across all image types, the rankings for the various measurement criteria appear to be stratified in a stable fashion with very few changes in rankings⁶, and those changes in rankings are typically only a one position change in ranking (i.e., in terms of effective compression ratio RLE ranks 5th on Digital Live Scan Slap but drops to 6th for Digital Live Scan Flat). The one exception is PNG for Digital Live Scan Rolled images where PNG jumps from 4th place ranking for all other image types to 1st place ranking.

⁶ When examining the ranking data in this section, it would be prudent to examine the actual values of the metric in question as a change in rank order may be the result of a very small change in the underlying metric.

4.5. Investigative Goal 5: Comparison of JPEG 2000 Implementation With Respect to Effective Compression Rate

With certain algorithms, performance of the algorithm can vary greatly by the specific implementation of that algorithm (such as different vendors producing compatible but different algorithms), or the architecture for which the algorithm was built (such as 32-bit vs. 64-bit). The fifth investigative goal of this study is to see if such implementation differences do indeed introduce significant operational differences in terms of effective compression ratio. Since the focus of this study is the JPEG 2000 algorithm, this examination was limited to lossless JPEG 2000 and does not include RLE, ICER or PNG.

4.5.1. Investigative Analysis 5

The median compression rate data in Table 11 shows that the various implementations of JPEG 2000 performed nearly identically. Pair-wise examination of the data using the Wilcoxon Signed Rank Test with an alpha of 0.05 however shows significant differences at values which appear to be identical. Further examination of the statistical processes used have lead the team to conclude that pairwise analysis of real numbers may not be ideal for comparisons of nearly identical real values where subtle but operationally insignificant differences can lead to statistically significant differences between two such lists of values. A further discussion of this effect is provided in Section 6 of this report.

Table 11 - Effective Compression Rates (Medians) by Implementation, Higher is Better

CODEC	All Data	Latent	Ink Card Scan				Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
OpenJPEG v.1.3, 32-bit	1.9550	2.3406	1.6978	1.5098	1.7607	1.7420	2.5555	2.9099	1.8950	2.7331
OpenJPEG v.1.3, 64-bit	1.9550	2.3406	1.6978	1.5098	1.7607	1.7420	2.5555	2.9099	1.8950	2.7331
OpenJPEG v.1.4, 32-bit	1.9551	2.3406	1.6978	1.5099	1.7608	1.7420	2.5555	2.9100	1.8950	2.7332
OpenJPEG v.1.4, 64-bit	1.9551	2.3406	1.6978	1.5099	1.7608	1.7420	2.5555	2.9100	1.8950	2.7332
JasPer 1.900.1 32-bit	1.9551	2.3407	1.6978	1.5099	1.7609	1.7420	2.5555	2.9099	1.8951	2.7331
JasPer 1.900.1 64-bit	1.9551	2.3407	1.6978	1.5099	1.7609	1.7420	2.5555	2.9099	1.8951	2.7331

Table 12 - Effective Compression by Implementation, Wilcoxon Signed Rank Test, at alpha = 0.05

Case	Comparison Pair	Image Type										
		CODEC 1	CODEC 2	All Data	Latent	Ink Card Scan			Digital Live Scan			
				All	All	All	Rolled	Flat	Slap	All	Rolled	Flat
1	OpenJPEG v.1.3, 32-bit	OpenJPEG v.1.3, 64-bit	0.3173	1.0000	0.3173	1.0000	0.3173	1.0000	1.0000	1.0000	1.0000	1.0000
2	OpenJPEG v.1.3, 32-bit	OpenJPEG v.1.4, 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	OpenJPEG v.1.3, 32-bit	OpenJPEG v.1.4, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	OpenJPEG v.1.3, 32-bit	JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	<0.0001	<0.0001
5	OpenJPEG v.1.3, 32-bit	JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	<0.0001	<0.0001
6	OpenJPEG v.1.3, 64-bit	OpenJPEG v.1.4, 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	OpenJPEG v.1.3, 64-bit	OpenJPEG v.1.4, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	OpenJPEG v.1.3, 64-bit	JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	<0.0001	<0.0001
9	OpenJPEG v.1.3, 64-bit	JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	<0.0001	<0.0001
10	OpenJPEG v.1.4, 32-bit	OpenJPEG v.1.4, 64-bit	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	OpenJPEG v.1.4, 32-bit	JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0012	<0.0001	<0.0001	<0.0001
12	OpenJPEG v.1.4, 32-bit	JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0012	<0.0001	<0.0001	<0.0001
13	OpenJPEG v.1.4, 64-bit	JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0012	<0.0001	<0.0001	<0.0001
14	OpenJPEG v.1.4, 64-bit	JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0012	<0.0001	<0.0001	<0.0001
15	JasPer 1.900.1 32-bit	JasPer 1.900.1 64-bit	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

4.5.2. Investigative Result 5

Examination of various implementations of JPEG 2000 in both 32-bit and 64-bit architectures yielded nearly identical results in effective compression rates therefore architectural differences are not expected to be a factor in the adoption of a particular implementation of lossless JPEG 2000 CODECs.

4.6. Investigative Goal 6: Comparison of JPEG 2000 Implementation With Respect to Compression Throughput

With certain algorithms, throughput of the algorithm can also vary greatly by the specific implementation of that algorithm (such as different vendors producing compatible but different algorithms), or the architecture for which the algorithm was built (such as 32-bit vs. 64-bit). The sixth investigative goal of this study is to see if such implementation differences do indeed introduce significant operational differences in terms of compression throughput. As with the previous section, since the focus of this study is the JPEG 2000 algorithm, this investigative goal was limited to lossless JPEG 2000 and does not include RLE, ICER or PNG.

4.6.1. Investigative Analysis 6

The median compression throughput data in Table 13 shows that there are differences evident between the various implementations of JPEG 2000. Pair-wise examination of the data using Wilcoxon Signed Rank Test with an alpha of 0.05 also confirmed significant differences between the OpenJPEG 32-bit v1.3 and 32-bit v1.4 implementations for all cases except Ink Scan Flat, Digital Live Scan All and Digital Live Scan Rolled.

Table 13 - Compression Throughput (Median Time, seconds) by Implementation, Lower is Better

CODEC	Image Type										
	All Data		Latent		Ink Card Scan			Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap	
OpenJPEG v.1.3, 32-bit	1.0506	1.0771	0.9126	0.7956	0.3762	3.8532	1.1811	1.1801	0.4614	3.2632	
OpenJPEG v.1.3, 64-bit	0.9940	1.0120	0.8658	0.7332	0.3515	3.7440	1.0980	1.0980	0.4266	3.1001	
OpenJPEG v.1.4, 32-bit	1.0757	1.0995	0.9282	0.7800	0.3732	4.0248	1.1775	1.1775	0.4527	3.3267	
OpenJPEG v.1.4, 64-bit	0.4700	0.4873	0.4042	0.3427	0.1641	1.7316	0.5130	0.5130	0.1904	1.4411	
JasPer 1.900.1 32-bit	0.4147	0.4289	0.3748	0.3234	0.1558	1.6068	0.4314	0.4314	0.1752	1.2050	
JasPer 1.900.1 64-bit	0.6708	0.7075	0.5928	0.4991	0.2434	2.5116	0.7152	0.7152	0.2846	2.0416	

Table 14 - Compression Throughput by Implementation, Wilcoxon Signed Rank Test, at alpha = 0.0033

Case	Comparison Pair	Image Type																		
		CODEC 1		CODEC 2		All Data		Latent		Ink Card Scan			Digital Live Scan							
		All	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap								
1	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.3, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.4, 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0758	<0.0001	0.4226	0.2151	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.4, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	OpenJPEG v.1.3, 32-bit vs JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	OpenJPEG v.1.3, 32-bit vs JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	OpenJPEG v.1.3, 64-bit vs OpenJPEG v.1.4, 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	OpenJPEG v.1.3, 64-bit vs OpenJPEG v.1.4, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	OpenJPEG v.1.3, 64-bit vs JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	OpenJPEG v.1.3, 64-bit vs JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	OpenJPEG v.1.4, 32-bit vs OpenJPEG v.1.4, 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	OpenJPEG v.1.4, 32-bit vs JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	OpenJPEG v.1.4, 32-bit vs JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	OpenJPEG v.1.4, 64-bit vs JasPer 1.900.1 32-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
14	OpenJPEG v.1.4, 64-bit vs JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
15	JasPer 1.900.1 32-bit vs JasPer 1.900.1 64-bit	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

4.6.2. Investigative Result 6

Examination of various implementations and architectures of JPEG 2000 yielded statistically significant differences in compression throughput between OpenJPEG 32-bit v1.3 and 32-bit v1.4 implementations in all but three cases: Ink Scan Flat, Digital Live Scan All and Digital Live Scan Rolled. Leading the pack is the highly optimized 32-bit implementation of the JasPer CODEC, followed by the 64-bit implementation of the OpenJPEG v.1.4 CODEC.

4.7. Investigative Goal 7: Comparison of JPEG 2000 Implementation With Respect to Decompression Throughput

As with compression throughput, the performance of algorithms can vary for the decompression process by implementation (i.e., same algorithm but different vendors or system architecture). The seventh investigative goal of this study is to see if such implementation differences do indeed introduce significant operational differences in terms of decompression throughput. Once again since the focus of this study is the JPEG 2000 algorithm, this investigative goal was limited to lossless JPEG 2000 and does not include RLE, ICER or PNG.

4.7.1. Investigative Analysis 7

The median decompression throughput data in Table 15 shows that there are differences evident between the various implementations of JPEG 2000. Pair-wise examination of the data using Wilcoxon Signed Rank Test with a significance level of 0.05 also confirmed significant differences for all of these cases but two (OpenJPEG 32-bit v1.3 vs. 32-bit v1.4 with Ink Card Scan Rolled, and OpenJPEG 32-bit v1.4 vs. JasPer 32-bit with Latent imagery).

Table 15 - Decompression Throughput (Median Time, seconds) by Implementation, Lower is Better

CODEC	Image Type									
	All Data	Latent	Ink Card Scan				Digital Live Scan			
	All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
OpenJPEG v.1.3, 32-bit	0.9203	0.9399	0.7877	0.6785	0.3245	3.4319	1.0434	1.0434	0.3958	2.9190
OpenJPEG v.1.3, 64-bit	0.9029	0.9206	0.7954	0.6707	0.3216	3.4006	0.9984	0.9984	0.3893	2.8150
OpenJPEG v.1.4, 32-bit	0.8578	0.8734	0.7174	0.6083	0.2964	3.1198	0.9494	0.9494	0.3482	2.6870
OpenJPEG v.1.4, 64-bit	0.3587	0.3749	0.3086	0.2617	0.1236	1.3258	0.3930	0.3930	0.1435	1.1019
JasPer 1.900.1 32-bit	0.3631	0.3635	0.3302	0.2871	0.1344	1.3726	0.3859	0.3859	0.1566	1.0729
JasPer 1.900.1 64-bit	0.5728	0.5874	0.4994	0.4320	0.2081	2.1527	0.6179	0.6179	0.2404	1.7299

Table 16 - Decompression Throughput, Wilcoxon Signed Rank Test, alpha of 0.05

Case	Comparison Pair	Image Type											
		CODEC 1	CODEC 2	All Data	Latent	Ink Card Scan			Digital Live Scan				
				All	All	All	Rolled	Flat	Slap	All	Rolled	Flat	Slap
1	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.3, 64-bit			<0.0001	<0.0001	<0.0001	0.4836	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.4, 32-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	OpenJPEG v.1.3, 32-bit vs OpenJPEG v.1.4, 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	OpenJPEG v.1.3, 32-bit vs JasPer 1.900.1 32-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	OpenJPEG v.1.3, 32-bit vs JasPer 1.900.1 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	OpenJPEG v.1.3, 64-bit vs OpenJPEG v.1.4, 32-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	OpenJPEG v.1.3, 64-bit vs OpenJPEG v.1.4, 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	OpenJPEG v.1.3, 64-bit vs JasPer 1.900.1 32-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	OpenJPEG v.1.3, 64-bit vs JasPer 1.900.1 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	OpenJPEG v.1.4, 32-bit vs OpenJPEG v.1.4, 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	OpenJPEG v.1.4, 32-bit vs JasPer 1.900.1 32-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	OpenJPEG v.1.4, 32-bit vs JasPer 1.900.1 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	OpenJPEG v.1.4, 64-bit vs JasPer 1.900.1 32-bit			<0.0001	0.2684	<0.0001	<0.0001	<0.0001	<0.0001	0.00029	<0.0001	<0.0001	<0.0001
14	OpenJPEG v.1.4, 64-bit vs JasPer 1.900.1 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
15	JasPer 1.900.1 32-bit vs JasPer 1.900.1 64-bit			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

4.7.2. Investigative Result 7

Examination of various implementations of JPEG 2000 in both 32-bit and 64-bit architectures yielded statistically significant differences in decompression throughput in all but two cases. These were OpenJPEG 32-bit (v1.3 and v1.4) on Ink Card Scan Rolled and Digital Live Scan Rolled, and OpenJPEG 32-bit v1.4 and JasPer 32-bit on Latent. Leading the pack is the 64-bit implementation of the OpenJPEG v.1.4 followed by the 32-bit implementation of the JasPer CODEC. It should also be noted that for the same algorithm version, the 32-bit compilation of that version provided lower decompression throughput than the 64-bit in almost all the cases, except with the JasPer implementation which favors the 64-bit platform.

4.8. Investigative Goal 8: Comparison of Implementation Complexity

Another factor that can impact the selection of an algorithm is the complexity of the code relative to the target platform being used. A true complexity analysis requires profiling of the algorithm with instrumented source code during typical execution. A rudimentary method of complexity analysis can be an estimation of complexity based on the lines of code in the source files that comprise the algorithm implementation. This complexity estimation is anecdotal and is not considered a good measure of complexity as a few lines of code may be executed at an exponential redundancy while many lines of code may execute in linear fashion. This section has been included nonetheless for discussion purposes.

4.8.1. Investigative Analysis 8

The source files of each respective algorithm were processed; the number of lines of code (LOC), the number of comments/empty line breaks, and the number of individual source files were all measured. The results of these measures are presented below in Table 17.

Table 17 - Implementation Complexity

CODEC	Gross Lines of Code (LOC)	Lines of Comments/Blanks	Net Lines of Code (LOC)	Number of Source Files
ICER	6883	991	5892	19
RLE used in BMP	39812	1713	38099	42
JasPer	42235	3857	38378	108
OpenJPG 1.3	70076	11601	58475	163
PNG (LIBPNG+ZLIB)	78893	9094	69799	116
OpenJPG 1.4	87676	14257	73419	179

4.8.2. Investigative Result 8

Based on the LOC analysis, the highly optimized ICER CODEC provides the lowest LOC count, and may therefore be the least complex. Based on the intended purpose of ICER being run on an autonomous space craft with limited computational resources, this finding would agree with that goal as ICER’s codebase appears to be a fraction of even the second least complex algorithm, RLE.

5. Conclusions

This experiment was conducted with eight primary goals:

1. Examine Effective Compression Rates by Algorithm
2. Examine Compression Throughput by Algorithm
3. Examine Decompression Throughput by Algorithm
4. Examine Effective Compression Rate by Impression Type
5. Compare JPEG 2000 Implementation With Respect to Effective Compression Rate
6. Compare JPEG 2000 Implementation With Respect to Compression Throughput
7. Compare JPEG 2000 Implementation With Respect to Decompression Throughput
8. Compare Implementation Complexity

Based on the results, it can be concluded that the wavelet-based algorithms (JPEG 2000 and ICER) provide far better effective compression rates than their non-wavelet-based counterparts (PNG and RLE using BMP) on mixed-image data consisting of Ink Card Scan, Digital Live Scan, and Latent. Given specific impression types, the PNG algorithm leads all others for the case of Digital Live Scan Rolled. It is hypothesized that the lower entropy of this particular image type may make these images more effectively compressible by PNG.

In terms of compression throughput, the non-wavelet-based algorithms lead the pack with the exception of the highly optimized JasPer 2000 (32-bit) which provides compression throughputs on par with the non-wavelet-based algorithms. With decompression however, PNG clearly leads the pack by a large margin. While PNG has many uses, it was primarily designed to be a format suitable for the transmission and display of raster images transmitted across networks and displayed in web browsers. This is a highly asymmetric use case where images are typically compressed once and viewed/decompressed many times. Optimization for such use cases, (compress once, view many times), as well as algorithm complexity, have provided PNG this edge in decompression throughput.

In examining algorithm performance with respect to the various image/impression types used in this study, the image type used did not appear to have a large impact on changing the rank order dramatically. Where a ranking shift was noted, this shift was typically only one rank position (i.e., an algorithm that ranked 6th best for Ink Card Scan Rolled imagery may have fared slightly better in the 5th place with another image type for a given measurement criteria). The one exception was PNG, which jumped from a 4th place ranking with All Data in terms of effective compression rate to 1st place for Digital Live Scan Rolled. Considering its fast decompression times, the PNG algorithm may be particularly well suited for Digital Live Scan Rolled images.

In studying the behavioral differences of various platforms or implementations of a given algorithm, examination of various implementations of JPEG 2000 in both 32-bit and 64-bit architectures yielded nearly identical results in terms of effective compression rates with some very small differences. While these small differences were statistically significant in almost all cases, they are not operationally relevant (for example, where OpenJPEG v.1.3 yields an effective rate of 1.9550:1 and OpenJPEG v.1.4 yields an effective rate of 1.9551:1). The various architectures did exhibit statistically significant differences in throughput but these differences were not consistent enough to make a generalized conclusion. For example, in almost all cases, 32-bit OpenJPEG implementation was slower than the 64-bit OpenJPEG implementation. For all those cases, the reverse was true with JasPer where the 32-bit version excelled.

Finally, based on a simple comparison of the number of lines of code that comprised each algorithm, the ICER CODEC is comprised of the fewest lines of code, and may perhaps be the least complex. The next two algorithms are RLE and JasPer.

Based on this study it can be generally concluded (with specific exceptions as noted) that the non-wavelet-based algorithms provide an edge in throughput while the wavelet-based algorithms provide an edge in effective compression rates.

6. Discussions and Future Work

Throughout this study, most of the selected measures of performance yielded real (floating point) numbers of very high precision. Comparison of such numbers in cases of very similar (but different) value may yield differences which can be operationally inconsequential, but would nonetheless be statistically significant. In this study, we attempted to mitigate this by conducting preliminary mathematical operations at six significant digit precision with the final stages of analysis at a precision of four. Such a strategy may only be effective in the scope of image sizes used in this study and generalization of it may call for adjustment of the rounding depending on the distribution of expected image sizes. This may be a ripe topic for future study where the very metrics of compression parameters can be explored in a framework relevant to the systems for which the operations are intended.

References

Publications and Reports

BRADLEY1	Bradley, Jonathan N., Brislawn, Christopher M., Hopper, Thomas. (1993). "FBI wavelet/scalar quantization standard for grayscale fingerprint image compression", SPIE Conference on <i>Visual Information Processing II</i> , eds. Huck, Friedrich O., Juday, Richard D. pp. 293-304.
BRADLEY2	Bradley, J. N., Brislawn, C. M. (1994). "The wavelet/scalar quantization compression standard for digital fingerprint images". <i>IEEE International Symposium on Circuits and Systems, 1994. ISCAS '94., 1994. Vol 3.</i> pp. 205-208.
BRISLAWN1	Brislawn, Christopher M. (1996). "Wavelet Scalar Quantization Compression Standard for Fingerprint Images". <i>Proceeding Conference on Signal Image Processing and Applications.</i>
BRISLAWN2	Brislawn, Christopher M. (2002). The FBI Fingerprint Image Compression Standard. http://www.c3.lanl.gov/~brislawn/FBI/FBI.html , updated 25 June 2001 (accessed 01/31/2014).
CHAMBERS	Chambers, John; William Cleveland, Beat Kleiner, and Paul Tukey (1983). <i>Graphical Methods for Data Analysis.</i> Wadsworth
CIA	"The World Factbook", https://www.cia.gov/library/publications/the-world-factbook/geos/us.html , Retrieved 2010-12-06.
CURRAN- EVERETT	Curran-Everett, Douglas (2000). "Multiple comparisons: philosophies and illustrations", <i>American Journal of Physiology - Regulatory, Integrative and Comparative Physiology</i> , 279, R1 – R8.
DOJ	"The Science of Fingerprints" [rev. December 1984], United States Department of Justice, Federal Bureau of Investigation, USGPO, ISBN 0-16-050541-0, Page 18.
EYPASCH	Eypasch, Ernst; Rolf Lefering, C K Kum, Hans Troidl (1995-09-02). "Probability of adverse events that have not yet occurred: a statistical reminder". <i>BMJ</i> 311 (7005): 619-620. PMID 7663258. PMC 2550668. http://www.bmj.com/cgi/content/full/311/7005/619 . Retrieved 2008-04-15.
FITZPATRICK	Fitzpatrick, M. et al. 1994, "WSQ Compression / Decompression Algorithm Test Report", IAI Annual Conference.
GALTON	Galton, F. (2005). <i>Finger prints.</i> Mineola, NY: Dover Publications. (Original work published 1892)
HANLEY	Hanley, JA; Lippman-Hand A (1983). "If nothing goes wrong, is everything alright?". <i>JAMA</i> 249 (13): 1743-5. PMID 6827763.
HOLLANDER	Hollander, M., Wolfe, D. A. (1999). <i>Non-Parametric Statistical Methods</i> , 2nd Ed, John Wiley & Sons:NY, 787 pages.
HOPPER	Hopper, T., Preston, F. (1992). "Compression of grey-scale fingerprint images", <i>Proceedings of Data Compression Conference, 1992. DCC '92.</i> pages 309-318.
JAIN	Jain, A., "Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features", <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , Vol. 29, No. 1, January 2007.
JOVANOIV	Jovanoviv, B. D., and P. S. Levy. A look at the Rule of Three. <i>The American Statistician</i> , Vol. 51, No. 2, May 1997, pp. 137-139.
KIELY1	Kiely, A. and Klimesh, M. "The ICER Progressive Wavelet Image Compressor," <i>The Interplanetary Network Progress Report 42-155, July–September 2003</i> , Jet Propulsion Laboratory, Pasadena, California, pp. 1–46, November 15, 2003. http://ipnpr.jpl.nasa.gov/progress report/42-155/155J.pdf
KIELY2	Kiely, A. and Klimesh, M. "Preliminary Image Compression Results from the Mars Exploration Rovers," <i>The Interplanetary Network Progress Report</i> , vol. 42-156, Jet Propulsion Laboratory, Pasadena, California, pp. 1–8, February 15, 2004. http://ipnpr.jpl.nasa.gov/progress report/42-156/156I.pdf
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). <i>A Technique for the Measurement of Attitudes</i> , Archives of Psychology 140, 55.
MTR	"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	National Institute of Standards and Technology. <i>Summary of NIST Patriot Act Recommendations</i> . Gaithersburg, MD. Retrieved January 4, 2007 from http://www.itl.nist.gov/iad/894.03/pact/NIST_PACT_REC.pdf
NIST2	"NIST Biometric Image Software". http://Fingerprint.nist.gov/NFIS/ . Retrieved 2011-01-12.
NISTIR7778	Orandi, S., Libert, J. M., Grantham, J. D., Ko, K., Wood, S.S., Wu, J. <i>Effects of JPEG 2000 Image Compression on 1000 ppi Fingerprint Imagery</i> , NIST Interagency Report 7778, National Institutes of Standards and Technology, Gaithersburg, MD. April 11, 2011, 72 pages.
NISTIR7781	Libert, J. M., Orandi, S., Grantham, J. D. <i>Comparison of the WSQ and JPEG 2000 Image Compression Algorithms on 500 ppi Fingerprint Imagery</i> , NIST Interagency Report 7781, National Institutes of Standards and Technology, Gaithersburg, MD March 2012, 56 pages.
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)". <i>Biometrika</i> 52 (3-4): 591–611. doi:10.1093/biomet/52.3-4.591. JSTOR 2333709MR205384.
WILCOXON	Wilcoxon, Frank (Dec 1945). " Individual comparisons by ranking methods ". <i>Biometrics Bulletin</i> 1 (6): 80–83.
WRIGHT	Wright, Paul, S. P-values for simultaneous inference. <i>Biometrics</i> , Vol. 48, No. 4 (Dec. 1992), pp. 1005-1013.
WU1	Wu, Jin Chu, Alvin F. Martin, and Raghu N. Kacker. Measures, uncertainties, and significance test in operational ROC analysis, <i>Journal of Research of the National Institute of Standards and Technology</i> , 116(1), 517-537, (2011).
WU2	Wu, Jin Chu. Studies of Operational Measurement of ROC Curve on Large Fingerprint Data Sets Using Two-Sample Bootstrap. NISTIR 7449, U.S. Department of Commerce, National Institute of Standards and Technology, September 2007, 25 pages.
WU3	Wu, Jin Chu. Operational Measures and Accuracies of ROC Curve on Large Fingerprint Data Sets. NISTIR 7495, U.S. Department of Commerce, National Institute of Standards and Technology, May 2008, 23 pages.

Standards

AN2011	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 2011.
ISO/IEC	ISO/IEC 15948:2004 – Information technology - Computer graphics and image processing -- Portable Network Graphics (PNG): Functional specification. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=29581 Retrieved 2012-01-11.
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.fbi biospecs.org/docs/WSQ_Gray-scale_Specification_Version_3_1.pdf . Retrieved 2010-01-11.

Appendix A. Dataset⁷ Makeup

For the Ink Card Scan portion of the tests, the study utilized fingerprint images based on the Base 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]. This Ink Card Scan data was 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 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 FBI certified software (Appendix F compliant) and stored in a non-compressed format to ensure no compression anomalies were introduced into the original set.

For the Digital Live Scan portion of the tests, the study again utilized actual operational data captured during normal enforcement activities with collection dates ranging from 01/04/2010 through 04/13/2010. The Digital Live Scan data was stored so as never to have been subject to lossy compression.

Where possible, the image sets were equally balanced by gender, finger, pattern class and hand. It should be noted that balancing equally does not follow the natural demographic behavior of the population such as gender (48 % males/52 % female [CIA]) or pattern class (65 % Loops, 30 % Whorls, 5 % Arches [DOJ]). The goal in having equal distributions was to avoid the potential statistical bias of very small subsamples. That is, all subsamples were equally important with respect to compression irrespective of their relative incidence in the population.

Demographic Make-up of Ink Card Scan Datasets

Ink Card Scan images used in this study consisted of 200 each of Rolled, Flat, and Slap.

Table 18 - Ink Card Scan Data classification by Impression Type

All Data				
Impression Type	Males	Females	Right	Left
Flat Single Finger	100	100	96	104
Rolled Single Finger	100	100	96	104
Four Finger Slaps	100	100	100	100

Table 19 - Ink Card Scan Pattern Classification for Single Finger Images by Impression Type

Data From Females (Single Finger)					Data From Males (Single Finger)				
Pattern Class	Flat	Rolled	Right	Left	Pattern Class	Flat	Rolled	Right	Left
Arch	34	34	30	38	Arch	34	34	34	34
Loop	33	33	32	34	Loop	33	33	32	34
Whorl	33	33	34	32	Whorl	33	33	34	32
Total	100	100	96	104	Total	100	100	100	100

⁷ This dataset was chosen to meet constraints required by the NISTIR 7778, but this does not lessen its applicability to this study.

Table 20 - Ink Card Scan Pattern Classification for Single Finger Images by Finger (Females)

Data From Females (Single Finger)											
Pattern Class	R. Thumb	R. Index	R. Middle	R. Ring	R. Little	L. Thumb	L. Index	L. Middle	L. Ring	L. Little	Total
Arch	4	4	3	4	0	3	5	4	3	4	34
Loop	3	3	4	3	3	4	3	3	4	3	33
Whorl	3	4	3	3	4	3	3	4	3	3	33
Total	10	11	10	10	7	10	11	11	10	10	100

Table 21 - Ink Card Scan Pattern Classification for Single Finger Images by Finger (Males)

Data From Males (Single Finger)											
Pattern Class	R. Thumb	R. Index	R. Middle	R. Ring	R. Little	L. Thumb	L. Index	L. Middle	L. Ring	L. Little	Total
Arch	4	3	3	4	3	3	4	3	3	4	34
Loop	3	3	4	3	3	4	3	3	4	3	33
Whorl	3	4	3	3	4	3	3	4	3	3	33
Total	10	100									

Make-up of the digital live Scan data sets

Digital live Scan images used in this study consisted of 200 Rolled,, 200 Flat, and 200 Four Finger Slap impressions.

Table 22 – Digital Live Scan Data Classification by Impression Type

All Data Records				
Impression Type	Males	Females	Right	Left
Flat Single Finger	100	100	100	100
Rolled Single Finger	100	100	100	100
Four Finger Slaps	100	100	96	104

Table 23 – Digital Live Scan Pattern Classification for Single Finger Images by Impression Type

Data From Females (Single Finger)					Data From Males (Single Finger)				
Pattern Class	Flat	Rolled	Right	Left	Pattern Class	Flat	Rolled	Right	Left
Arch	33	33	36	30	Arch	33	33	36	30
Loop	34	34	30	38	Loop	34	34	30	38
Whorl	33	33	34	32	Whorl	33	33	34	32
Total	100	100	100	100	Total	100	100	100	100

Table 24 – Digital Live Scan Pattern Classification for Single Finger Images by Finger (Females)

Data From Females (Single Finger)												
Pattern Class	R. Thumb	R. Index	R. Middle	R. Ring	R. Little	L. Thumb	L. Index	L. Middle	L. Ring	L. Little	Total	
Arch	4	4	4	3	3	3	3	3	3	3	33	
Loop	3	3	3	3	3	3	4	4	4	4	34	
Whorl	3	3	3	4	4	4	3	3	3	3	33	
Total	10	100										

Table 25 – Digital Live Scan Pattern Classification for Single Finger Images by Finger (Males)

Data From Males (Single Finger)											
Pattern Class	R. Thumb	R. Index	R. Middle	R. Ring	R. Little	L. Thumb	L. Index	L. Middle	L. Ring	L. Little	Total
Arch	4	4	4	3	3	3	3	3	3	3	33
Loop	3	3	3	3	3	3	4	4	4	4	34
Whorl	3	3	3	4	4	4	3	3	3	3	33
Total	10	100									

Data demographics

The fingerprint images used to compile the datasets as described above were taken from several subjects. The balancing of the samples used was based on the uniqueness of a single fingerprint and not individual subjects. As such, multiple, yet distinct, fingerprint impressions were taken from some subjects (i.e., some subjects contributed more than one finger).

Table 26 - Gender Breakdown for Data

Subjects by Gender and Race						
	Males	Females	White	Black	Hispanic	Asian
Ink Card Scan Dataset	72	17	38	47	3	1
Digital Live Scan Dataset	60	63	11	51	59	2
Combined Dataset (All)	132	80	49	98	62	3

Table 27 - Age Breakdown for Data

Subjects by Age								
	Under 18	18-24	25-29	30-34	35-39	40-44	45-49	50+
Ink Card Scan Dataset	5	54	18	3	1	4	3	1
Digital Live Scan Dataset	9	38	23	15	18	8	4	8
Combined Dataset (All)	14	92	41	18	19	12	7	9

Table 28 - Other Metadata: Height and Weight

Subjects by Height and Weight									
	<5'0"	5'0" – 5'5"	5'6" – 5'11"	6'0"+	<100 lbs	100 – 149 lbs	150 – 199 lbs	200 – 249 lbs	250+ lbs
Ink Card Scan Dataset	0	11	59	19	0	27	48	10	4
Digital Live Scan Dataset	4	45	63	11	1	45	56	17	4
Combined Dataset (All)	4	56	122	30	1	72	104	27	8

Table 29 - Other Metadata: Eye Color

Subjects by Eye Color					
	Brown	Black	Blue	Green	Hazel
Ink Card Scan Dataset	69	2	12	1	5
Digital Live Scan Dataset	111	1	2	6	3
Combined Dataset (All)	180	3	14	7	8

Table 30 - 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
Ink Card Scan Rolled	1016	1002	652	1718	1166	1165	643	2063	1170.8	1134.4	535.0	2568.6
Ink Card Scan Flat	602	592	444	843	785	801	497	1008	459.6	476.0	269.0	564.1
Ink Card Scan Slap	3192	3192	3045	3334	2009	2013	1744	2082	6264.4	6264.7	5339.3	6597.5
Digital Live Scan Rolled	1600	1600	1600	1600	1500	1500	1500	1500	2343.8	2343.8	2343.8	2343.8
Digital Live Scan Flat	687	674	515	825	1057	993	672	1500	725.1	652.5	377.9	1171.9
Digital Live Scan Slap	3200	3200	3200	3200	2000	2000	2000	2000	6250.0	6250.0	6250.0	6250.0

Appendix B. Equipment Used for Study

The equipment utilized for the processing of the image data with the algorithms described was comprised of a single PC customized specifically for the purposes of imaging software development, research, and testing. The specifications of this PC are as follows:

Model:	Dell Precision T7500
CPU:	2x Intel Xeon W5580 @ 3.20 GHz
Memory:	12.0 GB DDR3 (Registered ECC)
Storage:	2x OCZ Vertex 2 240GB SSD (RAID 1), 4x Western Digital 2TB HDD (RAID 0+1)
Operating System:	Microsoft Windows 7 64-bit Ultimate Edition

Appendix C. Examination of Entropy

Lossy compression may balance compressed file size against the discard of image information. That is, a lossy CODEC may achieve a desired compression ratio by varying the amount of information discarded, or achieve a specified level of fidelity to the original image by altering the size of the output compressed file, i.e., a lower compression ratio. Various CODECs provide varying degrees of user control over this balance, but in every case, the complexity of image content will set the stage for various optimization strategies to be executed by the compression algorithm in balancing information loss against compression ratio.

By contrast, lossless image compression, by definition, must preserve all information contained in the non-compressed original. Loss of information is not available to the lossless CODEC as part of its optimization scheme. Accordingly, the only option available to the lossless algorithm is to increase the compressed file size (i.e., reduce the compression ratio) in proportion to the complexity of the input non-compressed original.

In the present study, it was observed that the BMP-RLE CODEC performed reasonably well with the Digital Live Scan images, but poorly with the Ink Card Scan images, actually resulting in a compressed file size larger than that of the non-compressed original. Suspecting that the RLE effectiveness to be related directly to image entropy differences, we measured the entropy of both Ink Card Scan and Digital Live Scan for comparison.

Run-length encoding (RLE) is a very simple form of data compression in which runs of data (sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run. Hence, images containing large regions of homogeneous gray level may be encoded very efficiently using RLE. By contrast, highly textured images containing few runs are less efficiently encoded via RLE.

Entropy (E) provides a measure of the average gray level variability of an image. It is defined as:

$$E_I = -\sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (\text{C.1})$$

where n = number of gray levels in the image, i.e. 256; x_i = the value of the i^{th} gray level; $p(x_i)$ = the probability of occurrence of the i^{th} gray level in the image. Thus, for a single channel (8 bit) image having $p(x_i) = 1/256$ for every i , the maximum entropy is 8, or 8 bits. Accordingly, an image displaying a relatively flat (uniform) gray level histogram should yield an entropy very near the maximum value of 8 bits. An image having less uniform distribution of gray levels, such as an image displaying large areas homogeneous in gray level, would be expected to yield a lower entropy measure.

Procedure

Entropy was measured for each of the non-compressed source images used in the present study. This included 1000 ppi digital Scans acquired from: a standard inked 10-print fingerprint cards, and a Digital Live Scan device employing the principal of Frustrated Total Internal Reflection (FTIR).

In addition to entropy as defined in equation (1), we measured the ratio of entropy in the image area identified by a segmentation procedure as occupied by the background to that of the fingerprint. Hence, we have the Entropy Ratio defined as simply:

$$E_{\text{Ratio}} = \frac{E_{\text{background}}}{E_{\text{fingerprint}}} \quad (\text{C.2})$$

Results

Table 31 exhibits mean values of various entropy measures for Ink Card Scan data and those acquired using an FTIR Digital Live Scan device. The table resolves the mean values for each of the fingerprint impression types, rolled, flat, and 4-finger slap. Also included are measures of the proportion (percentage) of the image area identified via the segmentation procedure as containing the fingerprint.

Table 31 – Mean measurements of entropy and fingerprint region of images

CODEC		Entropy Image	Entropy Fingerprint	Entropy Background	Entropy Ratio	Prop Area Fingerprint
Rolled	Ink Card Scan	7.6	7.6	6.4	0.848	85.3
	Digital Live Scan	3.2	6.5	0.1	0.019	41.0
Flat	Ink Card Scan	7.4	7.4	6.0	0.811	63.7
	Digital Live Scan	5.7	6.9	1.6	0.227	66.3
Slap	Ink Card Scan	7.3	7.6	6.5	0.855	39.7
	Digital Live Scan	4.4	7.2	0.8	0.110	43.7

Entropy of the images tends to be higher for Ink Card Scan images in contrast to that of Digital Live Scan images. Entropy of fingerprint regions tends to be slightly lower for Digital Live Scan, but the main contrast is between entropy of the background regions, where that of Digital Live Scan is substantially lower than that of the Ink Card Scan.

Examination of the measures of the proportion of image area occupied by the fingerprint shows the greatest difference with the rolled impressions. This impression type, thus, is most likely to contain large areas of low-entropy background most efficiently compressed using the RLE algorithm. Such regions are more easily compressed by other algorithms as well yielding the highest effective compression ratios, i.e., smaller compressed file sizes.