NISTIR 7839

Examination of Downsampling Strategies for Converting 1000 ppi Fingerprint Imagery to 500 ppi

Shahram Orandi
John M. Libert
John D. Grantham
Margaret Lepley
Bruce Bandini
Kenneth Ko
Lindsay M. Petersen
Stephen S. Wood
Stephen G. Harvey

http://dx.doi.org/10.6028/NIST.IR.7839

National Institute of Standards and Technology
U.S. Department of Commerce

NISTIR 7839

Examination of Downsampling Strategies for Converting 1000 ppi Fingerprint Imagery to 500 ppi

Shahram Orandi
John M. Libert
Kenneth Ko
Stephen S. Wood
Stephen G. Harvey
Information Access Division - Image Group
Information Technology Laboratories

John D. Grantham
Systems Plus, Inc.

Margaret Lepley
Lindsay M. Petersen
MITRE Corporation

Bruce Bandini
Booz Allen Hamilton, Inc.

http://dx.doi.org/10.6028/NIST.IR.7839

January 2013

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 Federal Bureau of Investigation for all their support throughout this study.

The authors also give thanks to Vincent M. Stanford who provided early guidance on the creation and utilization of non-square filtering strategies in signal shaping which was used as a basis in the creation of the Gaussian filter utilized in this study.

In addition we are grateful for all the guidance, support and coordination provided by Michael D. Garris, whose help made this study possible.

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform evaluations. In no case does identification of any commercial product imply 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.

ABSTRACT

The Federal Bureau of Investigation (FBI) is increasing the sample rate at which it acquires and stores fingerprint imagery, and to accommodate the higher sample rate is replacing Wavelet Scalar Quantization (WSQ) with JPEG2000 as its standard compression algorithm. Currently the bulk of fingerprint data in operations is captured, processed and stored at 500 ppi (pixels per inch) using the WSQ compressed digital format. With the transition to 1000 ppi, some systems will unavoidably contain an overlap between 500 ppi and 1000 ppi operational pathways. This overlap may be a result of legacy infrastructure limitations, or some other financial or logistical reason. Additionally, there will still be a need to compare newly collected 1000 ppi images against legacy 500 ppi images, for both one-to-one and one-to-many scenarios. To create a bridge between legacy and modern data there needs to be a pathway for interoperability putting legacy and modern data on equal footing by converting one of the images to the same resolution as the other. Downsampling of the higher resolution 1000 ppi imagery to 500 ppi provides this pathway. The study compares several computational methods for downsampling of modern fingerprint images from 1000 ppi to 500 ppi. These treatments include pixel averaging, decimation, transcoding directly from the JPEG2000 codestream, Gaussian filtering and spectral truncation. Fidelity of downsampled 1000 ppi images relative to corresponding prints scanned natively at 500 ppi is evaluated via scoring by expert fingerprint examiners, an automated machine matcher, RMS differences and correlation of spectra. Gaussian low-pass filtering ($\sigma = 0.8475$, $r = 4$) combined with decimation emerged as the optimal downsampling strategy with respect to image fidelity when gauged by several different ranking strategies. Transcode methods involving selective decoding of JPEG2000 images so as to discard the highest resolution decomposition level yielded 500 ppi images that were not of optimal quality but were processed at considerably higher speed due to greater computational efficiency of such transcode algorithms making these methods attractive in a tradeoff between image fidelity vs. sheer computational throughput.

EXECUTIVE SUMMARY

The criminal justice community has traditionally exchanged and stored fingerprint imagery data at 500 pixels per inch[1] (ppi) or 19.7 pixels per millimeter (ppmm). The Wavelet Scalar Quantization (WSQ) fingerprint image compression algorithm is currently the standard for the lossy compression of 500 ppi fingerprint imagery and the WSQ Gray-Scale Fingerprint Image Compression Specification [WSQ] provides guidance based on an International Association for Identification (IAI) study [FITZPATRICK] to establish the acceptable amount of fidelity loss due to lossy compression in order for a WSQ encoder and decoder to meet FBI certifications. These certifications are designed to ensure adherence to the WSQ specification and thereby to ensure fidelity and admissibility in courts of law for images that have been processed by such encoders and decoders.

Modern biometric systems are now trending towards the capture, transfer and storage of fingerprint images at 1000 ppi or 39.4 ppmm which offers many benefits, notably greater fidelity and better representation of Level 3 features. The ANSI/NIST ITL [ANSI/NIST-ITL 1-2011] and ISO 19794 [ISO/IEC 15948:2004] standards require 1000 ppi image compression using the JPEG2000 algorithm rather than WSQ. The reason for this is due to design factors in the WSQ compression format that optimize it specifically for 500 ppi and making it unsuitable for supporting the finer detail present in 1000 ppi fingerprint imagery.

With the transition to 1000 ppi, some systems will unavoidably contain an overlap between 500 ppi and 1000 ppi operational pathways. This overlap may be a result of legacy infrastructure limitations, or some other financial or logistical reason. Additionally, there will still be a need to compare newly collected 1000 ppi images against legacy 500 ppi images, for both one-to-one and one-to-many scenarios. To create a bridge between legacy and modern data there needs to be a pathway for interoperability putting legacy and modern data on equal footing by converting one of the images to the same resolution as the other. Downsampling of the higher resolution 1000 ppi imagery to 500 ppi provides this pathway. Many well established computational methods exist for the downsampling of images.

The following study addresses the need for interoperability between the legacy 500 ppi data and modern 1000 ppi data by comparing several computational methods (treatments) for downsampling of modern fingerprint images from 1000 ppi to 500 ppi to allow for interoperability with legacy 500 ppi imagery. These treatments included pixel averaging, decimation, transcoding, Gaussian filtering and spectral truncation, which are measured using various techniques and the data analyzed to identify an optimal strategy.

To gauge the effectiveness of the downsampling methods several different fidelity measures were employed for each treatment including expert examiner assessment or machine matcher score assessment. The results of these fidelity measures were then fused using various techniques to provide an overall measures of how well the treatments worked. In order to provide a basis of measure for the downsampling treatments, a control set of native 500 ppi imagery that did not require any downsampling were utilized in the study in addition to the downsampled 1000 ppi images.

The experiment showed that the native 500 ppi control set of images yielded the best results versus 1000 ppi images that have been downsampled to 500ppi. Following the native 500 ppi control case were the Gaussian treatment cases.

The Gaussian treatment (σ = 0.8475, r = 4) emerged as the optimal downsampling strategy when gauged by several different ranking strategies including Ordinal Rank Summation (optimistic decimation), Mean Rank Summation (optimistic decimation) and Expert Examiner Winner-Take-All (which fuses optimistic and conservative decimation strategies). Where the Gaussian treatment (σ = 0.8475, r = 4) was not identified as optimal (following the control case), it was identified as second best (or tied as such) and was not a radical compromise in terms of the fidelity measures that failed to identify it as the best/top candidate.

Looking deeper into the various individual measures did uncover some surprises. In the case of expert examiner fidelity assessment, the downsampled images yielded better subjective quality than the native 500 ppi control case which indicates that some signal noise reduction may result in a positive impact in the perception of image quality by expert examiners.

With respect to objective examiner assessments of the images where the examiner was asked to make an identification decision between a downsampled image and the native 500 ppi version of that image, a statistically significant difference could not be established. This reinforces findings from NISTIR-7778 [NISTIR7778] that showed expert examiners could nearly universally detect degradation in fingerprint ridge structure as a result of lossy compression, but this effect did not appear to measurably impact their ability to conduct their identification tasks.

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

This is analogous to operating a motor vehicle with a "dirty windshield", where the driver's ability and experience mitigates any effect from occlusions large or small in their field of view.

An examination of automated (machine) matching of the downsampled images to a native 500 ppi gallery of fingerprint images showed that the 500 ppi native control case again outperformed all downsampling treatments, followed by several Gaussian treatments including (σ = 1.2730, r = 5, Odd-Even) using the optimistic decimation strategies and (σ = 1.1879, r = 5, Even-Odd) using the conservative decimation strategy. Based on the results of this study however, it appears that a strategy that optimized performance of the research matcher via very large Gaussian filters yielded a degradation in subjective fidelity assessment of the image by expert examiners. While the images favored by the matcher were not the worst in terms of expert examiner assessments, they did rank far lower by examiners. While it is possible to re-apply Gaussian filtering to an image, it is not possible to undo signal loss as a result of Gaussian filtering therefore a potential strategy to favor both the expert examiners as well as the matcher may involve applying one pass of Gaussian filtering in the downsampling process to provide some signal noise cleanup, and then later a second pass for some additional noise filtering just prior to processing of fingerprints for machine matching.

Examination of system throughput however points to Transcoded A and Transcoded B as the optimal solution in terms of operational efficiency and lower resource demands by eliminating several processing steps in the downsampling pathway. The elimination of processing steps by operating on the wavelet data directly helped avoid some of the intermediate steps needed for traditional downsampling and avoid the associated throughput penalty for these intermediate steps. Cursory studies show that these methods may be up to 10 times faster than the more computationally intensive Gaussian methods.

Finally, with regard to Gaussian filter sizes, it was shown that the automated matching favors Gaussian cases with larger sigma & radius values while other fidelity measures favor smaller sigma & radius value indicating that cases that are optimized for automated matching specifically may not be optimal for other fidelity measures including expert human perception.

VERSION HISTORY

Date	Activity
12/31/2012	Released final version 1.0

Table of Contents

List of Tables

LIST OF FIGURES

TERMS AND DEFINITIONS

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

Table 1 - Abbreviations

ANSI	American National Standards Institute
ANSI/NIST ITL	ANSI/NIST-ITL 1-2011, NIST Special Publication 500-290: Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information
bpp	Bits per pixel
FBI	Federal Bureau of Investigation
IAFIS	Integrated Automated Fingerprint Identification System
IAI	International Association for Identification
JPEG	Joint Photographic Experts Group - ISO/IEC committee developing standards for image compression
NBIS	NIST Biometric Image Software
NGI	Next Generation Identification
NIST	National Institute of Standards and Technology
ppi	Pixels per inch
ppmm	Pixels per millimeter
WSQ	The Wavelet Scalar Quantization algorithm for compression of fingerprint imagery

KEYWORDS

500 ppi; 1000 ppi; fingerprint imagery; downsampling; fingerprint compression; transcoding

1. Introduction and Investigative Goals

In partnership with the FBI, NIST launched an investigation, starting in November 2011, on the use of various downsampling strategies for converting 1000 ppi (39.4 ppmm) grayscale fingerprint images at 8 bpp to 500 ppi (19.7 ppmm). The 1000 ppi images were compressed with JPEG2000 [JPEG2K] per guidance provided in [MTR], and the 500 ppi were compressed with WSQ per FBI specifications. The methods were studied according to their impact on image quality and the ultimate utility of those images for machine based matching. This study encompassed the following objectives for downsampling comparison:

1. **Expert Examiner Subjective Assessment:** Investigate expert examiner opinion to gauge perceived degradation of 1000 ppi images as a result of downsampling to 500 ppi, compared to same image scanned natively at 500 ppi.
2. **Expert Examiner Objective Assessment:** Investigate impact on an expert examiner's ability to successfully make an ident/non-ident decision.
3. **Automated Matching Assessment:** Investigate effects on an automatic matcher by comparing match scores for pairs of downsampled 1000 ppi images versus native 500 ppi images.
4. **Signal Analysis and Assessment:** Investigate the effects of various downsampling strategies using various signal analysis methods.
5. **Optimal Strategy Determination:** Given the collected data, determine an optimal downsampling strategy for 1000 ppi fingerprint data.

1.1. Background

The criminal justice community has traditionally exchanged and stored fingerprint imagery data at 500 ppi, yet modern biometric systems are trending towards the transfer and storage of images at 1000 ppi. Transitioning to 1000 ppi images offers many benefits, notably greater fidelity to the original sample and better representation of Level 3 features. Both are favorable since they may increase probability of establishing a match/non-match by expert examiners or automated fingerprint matchers.

However, the transition to 1000 ppi systems will also unavoidably contain an overlap between 500 ppi and 1000 ppi operational pathways. Some stakeholders may simply continue to operate at 500 ppi due to technical, financial or logistical reasons. Additionally, there will still be a need to compare newly collected 1000 ppi images against legacy 500 ppi images, for both one-to-one and one-to-many scenarios.

There are two possible methods of interoperability between 500 ppi and 1000 ppi. The first method involves the downsampling (also known as subsampling) of 1000 ppi images to create a smaller 500 ppi image, while the second method would take a 500 ppi image and upsample it to 1000 ppi. The process of downsampling involves discarding data while upsampling would lead to the introduction of new data. Discarding data is more generally accepted for biometric data processing rather than the introduction of new data (through upsampling). The focus of this study is on the more generally acceptable practice of downsampling from 1000 ppi to 500 ppi to create a common ground for interoperability.

There are many recognized methods for downsampling of image data, each with their own set of advantages and disadvantages. This study focuses on four such methods: pixel averaging, decimation, transcoding, Gaussian filtering and spectral truncation. These methods are described in greater detail in section 3.1.

1.2. Key Drivers and Mandates

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 fingerprint analysis which began in 1965.

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. As part of a broader effort to enhance fingerprint system capacity, identification, and verification, this upgrade will help the FBI meet its mandate to:

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

Toward meeting these goals, the FBI seeks to guide the development of next generation friction ridge systems towards operation at 1000 ppi while at the same time ensuring interoperability with legacy 500 ppi systems, as well as maintaining fidelity and admissibility in courts of law.

The MITRE Corporation has developed an informative guidance that is widely recognized as the *de facto* standard guidance for utilizing JPEG2000 for the compression of 1000 ppi fingerprint imagery in MTR-04B0000022 [MTR]. This document provides a compression profile for 1000 ppi fingerprint imagery using JPEG2000 as well as guidance for two downsampling strategies that can be used to convert 1000 ppi JPEG2000 imagery to 500 ppi.

In support of the FBI, the National Institute of Standards and Technology (NIST) conducted a study to compare several downsampling strategies including the current *de facto* standard guidance [MTR] and measure the impact of the various strategies on operational and subjective criteria, and to formulate a basis with which normative downsampling guidance can be built.

2. Materials

2.1. Dataset

A total of 2039 images were selected for this study from an updated version of the NIST Special Database 27A (SD-27A) [SD27]. The SD-27A data set contains images that have been captured and stored in lossless JPEG-2000 format. Therefore these images have not been degraded in any way by lossy compression algorithms along their processing path and maintain complete fidelity to the original sample as recorded by the capture device. These images were selected from multiple subjects and were captured natively, referred to as "Native Scans", at 1000 ppi and 500 ppi without repositioning the original fingerprint card on the scanner platen, resulting in scanned images with approximately the same image content at multiple resolutions. The flatbed optical scanner used to capture these images was an Epson Perfection 4990 Photo scanner and was part of an FBI Appendix F certified commercially available solution. A thorough description of the data set used in this study can be found in Appendix A.

This set of 2039 1000 ppi fingerprint images was processed and stored in the lossy JPEG-2000 format per the guidance in [MTR]. The set was then subjected to 28 different downsampling treatments plus one Control Case (see Table 3). The rationale for converting the source images to the lossy JPEG-2000 format per the [MTR] guidance is because this case most closely resembles the operational Use Case of concern where a 1000 ppi image has been captured and stored at origination point per [MTR], and later needs to be processed for a legacy pathway. Other operational Use Cases exist such as a 1000 ppi scanner capturing an image at the collection station, saving it in uncompressed format, and then immediately downsampling it to 500 ppi at the collection station prior to processing at the site. These other Use Cases are considered outside the scope and were not examined since the study focus is on 1000 ppi images already stored in JPEG-2000 format. A full list of these Use Cases is provided below in Table 2.

Table 2 - Typical Use Cases that May Require Transcoding

Source Format	Destination Format	Notes
Lossless 1000 ppi	**Lossless 500 ppi**	**Not examined in this study.** Not an anticipated operational Use Case.
Lossless 1000 ppi	**WSQ lossy 500 ppi**	**Not examined in this study.** Useful for example for backwards compatibility where a 1000 ppi-only scanner would capture, downsample and provide images to a 500 ppi application/processing layer.
JPEG-2000 Lossy 1000 ppi	**Lossless 500 ppi**	**Not examined in this study.** Useful for example with on-site processing of 1000 ppi images in biometric systems that can only vectorize fingerprint images at 500 ppi and there is a desire to maintain operational benefit from the 500 ppi image without incurring any further degradation due to re-compression at 500 ppi after downsampling.
JPEG-2000 Lossy 1000 ppi	**WSQ lossy 500 ppi**	**The focus of this study.** Expected operational Use Case where fingerprints compressed according to [MTR] would be downsampled and submitted through a pathway designed for production WSQ 500 ppi imagery.

Of the 28 downsampling treatments, 24 required decimation for geometric dimension correction after processing (see Section 5.7). Rather than selecting a particular method of decimation, all four possible combinations of decimation were applied to those 24 cases, creating 96 sub-treatments for the cases requiring decimation. All downsampling treatments included WSQ compression[2] as a final output step. Each case was then paired with fingerprint images originally scanned at 500 ppi and WSQ compressed, referred to as "Native 500 ppi Scans". The Control Case contained two copies of the same 500 ppi native scan image which were both compressed with WSQ.

[2] R = 0.55 used as the WSQ compression rate control parameter.

Table 3 - Treatments Applied to Test Images

Treatment Number	Treatment Name	Decimation Sub-treatment(s)			
1	Native 500 ppi (Control Case)	None	-	-	-
2	Averaging	None	-	-	-
3	Decimation	Even-Even	Even-Odd	Odd-Even	Odd-Odd
4	NIST Spectral Truncation	None	-	-	-
5	Transcoded A	None	-	-	-
6	Transcoded B	None	-	-	-
7	Gaussian σ = 0.3369, r = 2	Even-Even	Even-Odd	Odd-Even	Odd-Odd
8	Gaussian σ = 0.4220, r = 2	Even-Even	Even-Odd	Odd-Even	Odd-Odd
9	Gaussian σ = 0.5071, r = 3	Even-Even	Even-Odd	Odd-Even	Odd-Odd
10	Gaussian σ = 0.5642, r = 3	Even-Even	Even-Odd	Odd-Even	Odd-Odd
11	Gaussian σ = 0.5922, r = 3	Even-Even	Even-Odd	Odd-Even	Odd-Odd
12	Gaussian σ = 0.6773, r = 3	Even-Even	Even-Odd	Odd-Even	Odd-Odd
13	Gaussian σ = 0.7624, r = 3	Even-Even	Even-Odd	Odd-Even	Odd-Odd
14	Gaussian σ = 0.8475, r = 4	Even-Even	Even-Odd	Odd-Even	Odd-Odd
15	Gaussian σ = 0.9326, r = 4	Even-Even	Even-Odd	Odd-Even	Odd-Odd
16	Gaussian σ = 1.0177, r = 4	Even-Even	Even-Odd	Odd-Even	Odd-Odd
17	Gaussian σ = 1.1028, r = 4	Even-Even	Even-Odd	Odd-Even	Odd-Odd
18	Gaussian σ = 1.1879, r = 5	Even-Even	Even-Odd	Odd-Even	Odd-Odd
19	Gaussian σ = 1.2730, r = 5	Even-Even	Even-Odd	Odd-Even	Odd-Odd
20	Gaussian σ = 1.3581, r = 5	Even-Even	Even-Odd	Odd-Even	Odd-Odd
21	Gaussian σ = 1.4432, r = 5	Even-Even	Even-Odd	Odd-Even	Odd-Odd
22	Gaussian σ = 1.5283, r = 6	Even-Even	Even-Odd	Odd-Even	Odd-Odd
23	Gaussian σ = 1.6134, r = 6	Even-Even	Even-Odd	Odd-Even	Odd-Odd
24	Gaussian σ = 1.6985, r = 6	Even-Even	Even-Odd	Odd-Even	Odd-Odd
25	Gaussian σ = 1.7836, r = 6	Even-Even	Even-Odd	Odd-Even	Odd-Odd
26	Gaussian σ = 1.8687, r = 7	Even-Even	Even-Odd	Odd-Even	Odd-Odd
27	Gaussian σ = 1.9538, r = 7	Even-Even	Even-Odd	Odd-Even	Odd-Odd
28	Gaussian σ = 2.0389, r = 7	Even-Even	Even-Odd	Odd-Even	Odd-Odd
29	Gaussian σ = 2.1240, r = 7	Even-Even	Even-Odd	Odd-Even	Odd-Odd

The image pairs (consisting of a native and downsampled image) from each treatment were then assessed using 4 different comparison methods to measure degradation as the result of the various treatments. These methods are listed below in Table 4 and are explained in detail in Section 3.1. Note that not all treatments utilized the same number of images. The automated matcher testing method utilized the full set of 2039 images, while the other methods utilized a subset of 200 images. This is because anecdotal evidence suggests the automated matcher behaves in a less predictable way for match score distribution as compared to the other metrics, thus requiring more images to enhance statistical analysis of the resulting distribution. The 200 sets of images selected for the other tests utilized the same set of 200 for all the comparison methods (there was no reselection per method) to allow for apples-to-apples comparison of the treatments.

Table 4 - Comparison Methods

Comparison Number	Comparison Method	Number of Image Pairs Used
1	Automated Matcher testing	2039
2	SIVV RMSD measurement	200
3	SIVV Correlation analysis	200
4	Latent Examiner analysis	200

2.2. Fingerprint Examiners

This study utilized 15 paid professional fingerprint examiners to look at fingerprint image pairs in their professional capacity and to render their professional judgment much as they do in their normal professional activities.

There was no attempt to evaluate the examiner's accuracy or proficiency for this study. There was also no attempt to identify or maintain the identity of the individual examiners utilized in this study.

The examiners had anywhere from three years to well over thirty years of experience in fingerprint examination. The examiners were recruited from various federal, state, local and commercial entities from across the United States.

A pre-requisite for examiners being selected to participate in this study was that they be trained *latent-print-examiners (LPEs)*. This requirement ensures that the examiners have received the most rigorous training possible. All 15 examiners participating in this study have earned the IAI's latent examiner certification (and are referred to as *certified latent print examiners* or *CLPEs*). Selection of examiners for this study was not meant to bias the results by utilizing latent-only examiners and all of the examiners selected for this study perform 10-print case work in addition to latent case work as part of their regular professional duties. While not the case in this study, it should be noted that some of the LPE's or CLPE's working with larger agencies may be assigned latent only case work with almost no 10-print case work.

The study provided three examiner workstations. The examiners were paid professionals and were allowed to work on the study according to their ability and timing. Each examiner looked at a random subset of the image pairs in the study and no one examiner observed all image pairs. No examiner looked at the same image pair more than once. Overall, each image pair was examined by exactly three examiners, but the selection of which three examiners viewed a given pair was done at random.

2.3. Equipment

The equipment used for the expert fingerprint examiner portion of this study consisted of the following (equipment used for other metrics is described in Appendix B):

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

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

3 x 24 inch Monitors: The IPS-panel monitors are connected via DVI-D and calibrated (see below) for optimal accuracy and consistency. The monitors were operated at their native resolution of 1920x1200, yielding a dpi measurement of approximately 94.3ppi. The zoom functionality in software allowed the examiners to zoom in and out of the image, and view them in the range of approximately 10x to 50x given these specific monitors.

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

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

[3] The FIXT image comparison software was developed for NISTIR-7778 [NISTIR7778] and has been reused in this study. More information on FIXT can be found in the methodology section of NISTIR-7778.

3. Methods

3.1. Downsampling Treatments

3.1.1. Decimation

This treatment utilizes an algorithm that eliminates entire rows and columns, resulting in an image that is half the resolution of the original. Specifically, the row and column removal strategy can be one of four possible combinations as shown below in Table 5. With odd-sized images (where there are an odd number of rows and/or columns), and a decimation strategy that eliminates odd rows and/or columns, the algorithm copies over the last odd row and/or column as a special condition.

The algorithm utilizes the most common image coordinate system, with the origin, (0, 0), placed at the upper-left corner of the image to determine whether a row or column is even or odd. The first row and column, with a coordinate value of 0, are considered even, and the second row and column with the coordinate value of 1 is considered odd. As such, a decimation strategy that retains even numbered rows and columns would retain rows and columns numbered 0, 2, 4, 6, 8, and so on in the new image resulting from the decimation process. A visual example of decimation strategy 1 from Table 5 is shown in Figure 1 below.

Table 5 - Decimation Pathways

Decimation Strategy	Removal Process
1	Retain Even Columns, and Even Rows
2	Retain Even Columns, and Odd Rows
3	Retain Odd Columns, and Even Rows
4	Retain Odd Columns, and Odd Rows

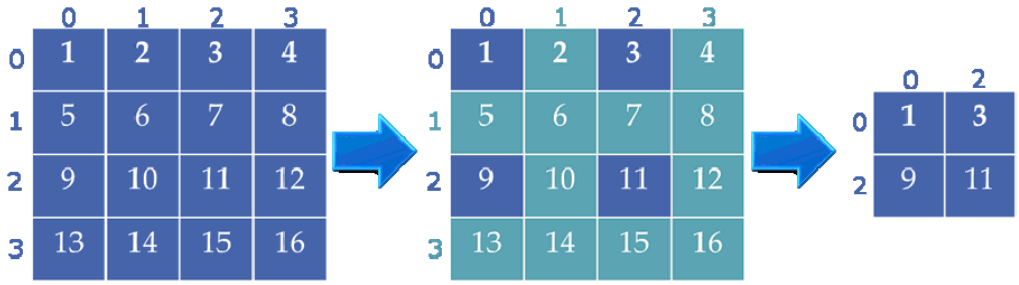


Figure 1 - Decimation

While decimation provides a simple and fast method of downsampling an image, the filter can be unpredictable in that it does not deal with frequency information in an elegant fashion and in certain cases, it can effectively fold high frequency image information into the lower frequency ranges. An example of such an undesirable effect is an image that consists of alternating white and black lines of one pixel in height and N pixels in width (Figure 2). In this extreme example, removal of odd rows and columns results in an image that is half of the resolution of the original but completely solid black. In effect, the high frequency pattern has been replaced with a lower frequency pattern.

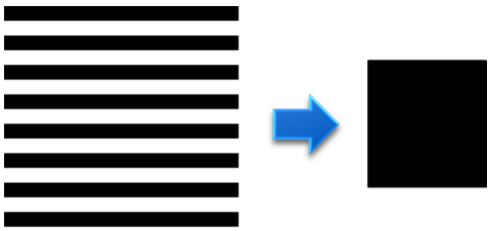


Figure 2 - Example of Undesirable Result with Decimation

It is expected that each of the four methods should perform equally as the others when applied to the fingerprint imagery selected for this study as the random nature of the data set was not expected to have dominance or bias towards odd or even decimation strategies. While no bias is expected in the data, for the purposes of completeness, all four decimation strategies from Table 5 were examined in this study.

3.1.2. Averaging

This treatment utilizes an algorithm that derives an intensity value for a single pixel value by summing the four neighboring pixels' intensity values and dividing the result by four to obtain an average intensity value representative of the four neighboring pixels. This averaged value will then replace the four adjacent pixels as one new pixel with the average as its intensity value. The result of this operation is an image that is one half the resolution of the original image. With odd-sized images (odd number of rows or columns), the algorithm simply copied over the last odd row and/or column as a special condition.

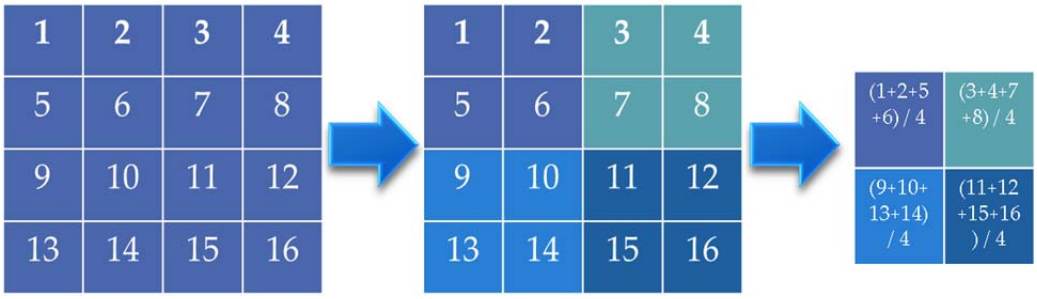


Figure 3 - Pixel Averaging Strategy

3.1.3. Transcoding

Transcoding typically refers to the act of making a change in the representation of compressed data by operating on the compressed data without going through a cycle of decompression and recompression.

Wavelet based compression algorithms such as JPEG2000 and WSQ generate a data stream of wavelet coefficients that represents both location and frequency information for discrete components that can be reassembled to create a representation of the original image upon decompression. The information the wavelet streams contain as well as the organization of the stream allows for operations on the stream that can effectively extract/filter a portion of the stream containing a desired portion of the spatial frequency spectrum. In particular, a wavelet transform contains a decimated low-pass filtered[4] version of the image which provides a representation of the image at half-resolution, as well as additional high-pass information necessary for a full resolution reconstruction.

For the purposes of this study, we utilized guidance provided by [MTR] that describes two transcoding strategies. The first method, referred to as Transcoded A operates by filtering an image compressed using guidance in [MTR] to decompress the low-pass 500 ppi portion of the frequency spectrum. Implementation of this method amounts to specifying to the JPEG2000 decoder that the highest resolution decomposition layer of the codestream is to be discarded. A parameter in the command line of the CODEC is all that is required to yield a 500 ppi image upon decoding the JPEG2000 codestream. The resulting 500 ppi decompressed image is then recompressed using WSQ. While this may not follow the literal definition of transcoding (where the image is operated on as a compressed stream rather than going through a decompression cycle), it does utilize aspects of the compression stream to filter frequency information resulting in an image downsampled to 500 ppi.

The second method referred to as Transcoded B operates in the more traditional sense of transcoding in that the compressed stream is operated on without fully decompressing the stream. The wavelet stream is filtered for the desired frequency and spatial information and

[4] Both JPEG2000 and WSQ use a low-pass filter using 9 coefficients decimated in an even-even fashion. The low-pass filter with approximate kernel weights of (0.03, -0.02, -0.08, 0.27, 0.60, 0.27, -0.08, -0.02, 0.03) in a single dimension is applied both horizontally and vertically. At the edges of the image where the kernel extends outside the image region, the kernel is instead reflected back from the border to reuse image pixels near the border. [The decimation offset is adjustable in JPEG2000, but the default even-even offset was used for all files in this experiment.]

converted to floating point array at 500 ppi that is compatible with the WSQ algorithm. Transcoded B's goal is to combine the JPEG2000 reduced resolution decoding with the WSQ encoding in a fashion that eliminates unnecessary loss of precision. Most of the JPEG2000 processing is performed in floating point[5], with a final conversion of the image pixels to only 8-bit accuracy for the purposes of output format expectations. Likewise, WSQ reads 8-bit image data, but immediately converts it to floating point before continuing processing. This conversion down to 8-bits causes numeric loss, which can be avoided by bypassing the down and up conversion steps, and moving floating point data directly from JPEG2000 machinery into WSQ floating point processing.

Figure 4 shows a typical JPEG2000 reduced resolution decode on the top, moving from left to right, following the blue arrows. On the bottom of the figure is the typical WSQ encode process, again following the blue arrows. The full path shown by the blue arrows is Transcoded A. Transcoded B eliminates the shaded area at the right, and instead follows the green arrow, directly between the two codecs.

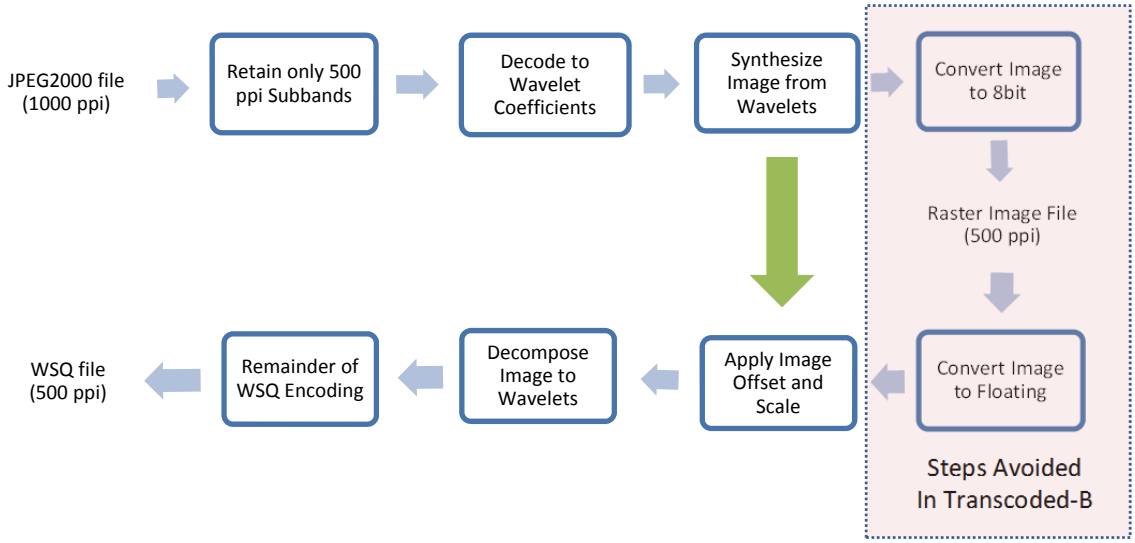


Figure 4 - Comparison of the Transcoded A and Transcoded B Methods

In addition to the significant speed-up provided by eliminating decode processing for the extra 1000 ppi sub bands, which is shared by both Transcoded A and Transcoded B methods, Transcoded B avoids a costly intermediate file save when efficiently implemented[6].

3.1.4. Gaussian Filtering with Decimation

Gaussian filtering or blurring is applied to an image prior to decimation in order to reduce the expression of frequencies sampled at a higher rate that would not have been present were the image formed via sampling at a lower rate. Also, termed "low-pass" filtering, the Gaussian is applied to reduce the effects of "aliasing" where in the decimated image the under-sampled high frequency components appear "aliased" as low frequency components. A common example of such is the moiré effect. Hence, prior to resampling the images at one-half the sample rate, we apply a Gaussian filter to the image in order to blur the image, which is equivalent to attenuating the high frequency content.

The Gaussian filter in two dimensions consists of a square array of weights, referred to as a Gaussian kernel, computed according to the expression

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (1)$$

where x and y are respectively distances in pixels from the center or origin, often referred to as the radius of the filter, and σ the standard deviation of the Gaussian. Implemented as a convolution operation on the intensity image, i.e., in the spatial domain, the filter array is centered on each pixel of the image and the weights multiplied by corresponding image pixels of the neighborhood of the image pixel being processed.

[5] Actual JPEG2000 implementations often use fixed point representations, with up to 32-bits of accuracy available.
[6] For this test, an intermediate output file of 32 bit signed integer data was generated by JPEG2000 and consumed by a customized version of WSQ.

The weighted sum resulting from this application of the filter is output to a new image. The Gaussian kernel is centered on the next pixel; the weighted summation again performed, and this continues over all rows of the image. At the edges of the image where the kernel extends outside the image region, the missing kernel values are filled in by "mirroring", or duplicating, the edge pixels.

Normally, it might be considered possible to design a Gaussian filter specifically to the alteration of sample rate desired. In practice, however, the specific effects of JPEG2000 and WSQ compression, their interactions with each other and with the original and output sample rates were considered difficult to predict or to control. Hence, we chose to conduct empirical tests employing a variety of filters in order to select the one or several best for the Use Case under consideration as described previously. Thus, the Gaussian filtering was repeated using each of the filters having parameters listed in Table 6.

Table 6 - Gaussian Filter Parameters Selected for This Study

Gaussian Strategy	Sigma	Radius (r)	Width
1	0.3369	2	5
2	0.4220	2	5
3	0.5071	3	7
4	0.5642	3	7
5	0.5922	3	7
6	0.6773	3	7
7	0.7624	3	7
8	0.8475	4	9
9	0.9326	4	9
10	1.0177	4	9
11	1.1028	4	9
12	1.1879	5	11
13	1.2730	5	11
14	1.3581	5	11
15	1.4432	5	11
16	1.5283	6	13
17	1.6134	6	13
18	1.6985	6	13
19	1.7836	6	13
20	1.8687	7	15
21	1.9538	7	15
22	2.0389	7	15
23	2.1240	7	15

Note relative to Table 6, the total width of the Gaussian kernels are computed as 2*radius + 1. The values of *x* and *y* inserted into equation (1) extend from *-r* to *+r*.

Following each application of the Gaussian filters, the filtered images are reduced to one-half the sample rate by each of the four decimation schemes as described in section 3.1.1. The C# computer code used to apply both the Gaussian filter is included in Appendix C.

3.1.5. Spectral Truncation

The discrete Fourier transform (DFT) converts a grayscale image of the spatial domain to a representation in the frequency domain. Barring any modification of the frequency space representation of an image, the inverse transform of the frequency representation reproduces the original grayscale image without loss. The grayscale image, *f(x,y)*, and its discrete Fourier transform, *F(u,v)*, are related by the following expressions:

$$F(u,v) = \frac{1}{NM}\sum_{x=1}^{N-1}\sum_{y=1}^{M-1} f(x,y)e^{-2\pi j(xu/N+yv/M)}$$

and **(2)**

$$f(x,y) = \sum_{x=0}^{N-1}\sum_{y=0}^{M-1} F(u,v)e^{2\pi j(xu/N+yv/M)}$$

where (*x*,*y*) and (*u*,*v*) are respectively image pixel coordinates and DFT wave numbers, *N* and *M* are width and length of the image and its DFT and j = $\sqrt{-1}$ ... Notably, an image and its DFT are simply alternate representations of the same information; having the one enables computation of the other without loss.

The DFT of an image of size *N* x *M* represents all frequencies present in an image sampled at 1000 ppi to the Nyquist frequency limit of 500 transitions per inch, scaled as symmetrical *N*/2 x *M*/2 *matrices*. It follows that the DFT of the image sampled at 500 ppi of size *N*/2 x *M*/2 should then consist of half the frequency range of the higher sample rate or *N*/4 x *M*/4. From this, we realize that a 500 ppi image may be derived from a 1000 ppi image by extracting from the DFT of the 1000 ppi image those frequency components which would be present in the image sampled at the lower rate. Applying the inverse transform to the rectangular region extracted from the DFT of the 1000 ppi yields a 500 ppi image, thus downsampled by spectral truncation. The mechanics of the procedure are outlined and demonstrated in the following discussion.

Figure 5 depicts a fingerprint impression scanned at 1000 ppi. We wish to downsample this image so as to derive an image sampled at 500 ppi. The Fourier transform of the image is shown in Figure 6. As has been discussed elsewhere, we might downsample the image by various other means. However, we can accomplish the adjustment in the sample rate of the image by excluding frequencies that would not be present at the lower sample rate. Noting the Fourier transform image (scaled logarithmically for display) the magnitudes of frequencies from the zeroth at the image center (the so-called direct current or DC level) to limits of one-half the 1000 ppi sample rate in positive directions to the right and up and negative to the left and below the DC value centered at (0,0). The rectangle inscribed on the interior of the DFT marks the portion that is extracted for inversion to create the 500 ppi downsampled image.

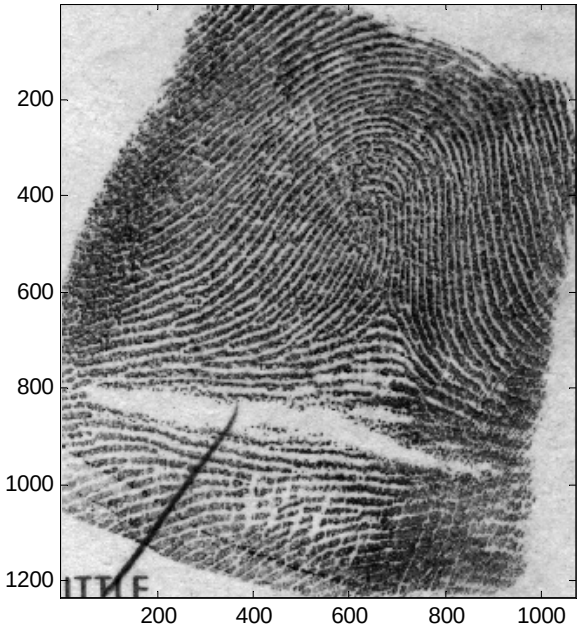


Figure 5 - Example of an inked, rolled fingerprint impression digitized at a sample rate of 1000 ppi.

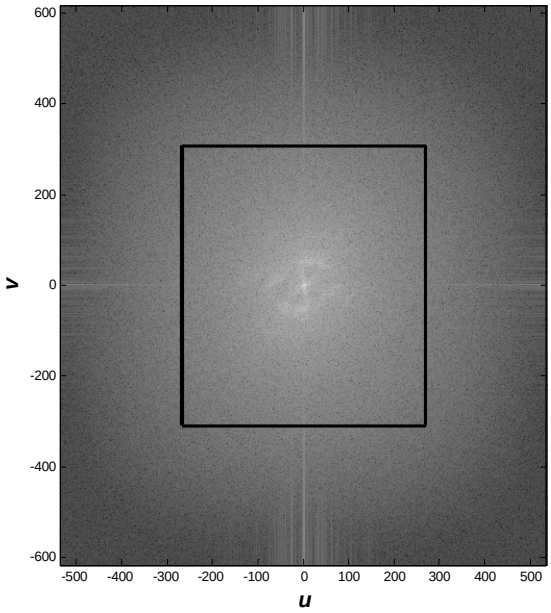


Figure 6 - The DFT of the grayscale image depicted in Figure 5 displayed as an image. The inscribed rectangle marks the portion of the spectrum to be extracted for construction of the 500 ppi image. (Note that the four quadrants of the spectrum have been translated diagonally so as to center the zero-frequency term, and the values have been scaled by the log function to enable display of the spectrum otherwise overwhelmed by the high magnitude of the DC value.)

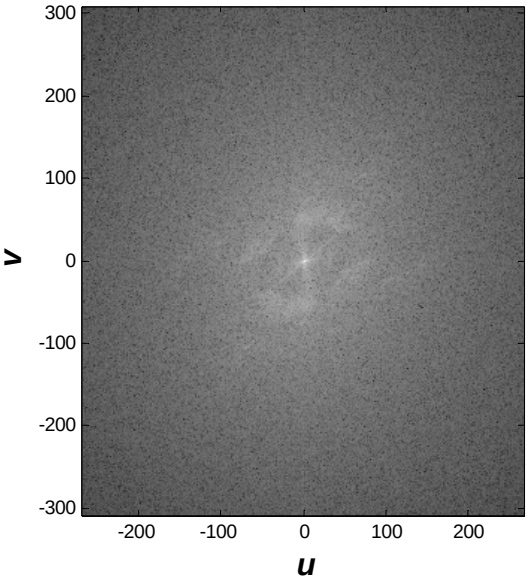


Figure 7 - The extracted portion of the DFT exhibited in Figure 6 to be inverted to form the 500 ppi downsampled image shown in Figure 8.

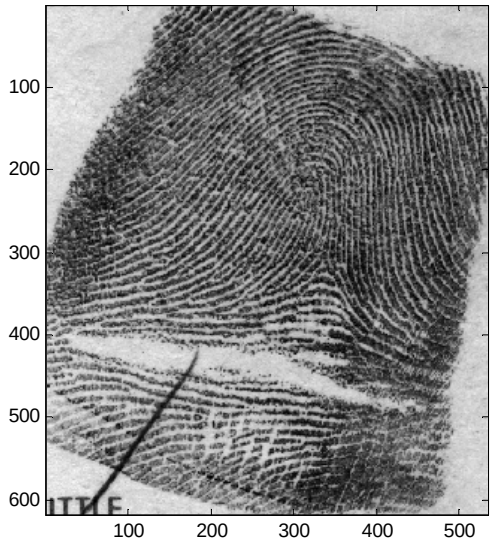


Figure 8 - The 500 ppi image formed from inverting the truncated spectrum (Figure 7) of the fingerprint originally sampled at 1000 ppi.

Several operational points should be noted relative to the spectral truncation method. First the complex valued DFT must be rearranged such that the four diagonally opposed quadrants are shifted across the center point such that the zero-frequency (dc) term is centered. Second, from this geometrically transformed DFT is extracted a rectangular region centered on the dc term having dimensions one half those of the 1000 ppi image and its DFT. Third, prior to application of the inverse transform, the extracted portion of the DFT must be returned to the conventional DFT geometry by reversing the diagonal translation of the quadrants. Fourth, the inverse DFT is applied yielding a 500 ppi image. Finally, it is noted that the grayscale of the resulting image is expanded relative to that of the 1000 ppi source image. Hence, the grayscale of the downsampled image must be rescaled to the intensity range of the 1000 ppi source image. This may be done by applying the rescaling function of equation (3) to the pixel values of the downsampled image.

$$I_{rescaled} = \left(\frac{I - I_{\min}}{I_{\max} - I_{\min}} \right) \left(I_{\max}^{source} - I_{\min}^{source} \right) \qquad (3)$$

where I is the downsampled image array to be rescaled, $I_{\min}$ and $I_{\max}$ are the lower and upper limits of the pixel values in I, $I_{\max}^{source}$ and $I_{\min}^{source}$ are the upper and lower limits of the grayscale of the 1000 ppi source image.

3.2. Metrics

3.2.1. Expert Examiners

Each of the native and downsampled images for all treatments/sub-treatments was shown to exactly three examiners. The image pairs were queued on each examiner's workstation and their presentation order was shuffled randomly on each of the three workstations. Each image pair was presented to 3 different examiners, and each image pair was viewed by each examiner only once.

The examiners were first asked to make an identification decision on the pair of images being shown to them. The identification decision had 3 possible choices: Yes, No and Inconclusive. The three choices were then post-processed using ground-truth data into the coded responses provided below in Table 7. Observation codes are in ascending order indicating a progressively greater negative impact on the identification process, with code 1 representing the right decision answer and code 3 representing the incorrect decision answer given the ground truth.

Table 7 - Identification Decision Coding

Observation Code	Description
1	The examiner answered the identification question correctly based on ground-truth information on the image pair shown to the examiner.
2	The examiner answered the identification question with an "Unknown" response indicating that based on the images shown, the examiner cannot with confidence note an identification or non-identification.
3	The examiner answered the identification question incorrectly based on the ground-truth information. For example, if the image pair shown represented a matched pair, they incorrectly indicated a non-identification. Similarly, if the image pair was a non-identification.

The post-processed coded responses from each of the 3 examiners were then concatenated so that a response of 1, 3 and 2 becomes 132, for example. This concatenated response was then sorted to list the triplet of values in increasing order, so 132 becomes 123. This process yields a concatenated response code that establishes a natural cost function as shown below in Table 8.

Table 8 - Sorted Examiner Response Code Classification

Responses from Observers	Classification
111 112 113	(BEST) These cases represent either a unanimous or majority ruling of observation code 1 from Table 9 indicating a correct identification or non-identification decision.
122 123 222 223	These cases represent either a unanimous or majority ruling of observation code 2 from Table 9. This case also includes the special split-decision case of 1, 2, 3. These cases are considered borderline cases and are biased towards the examiners being Inconclusive in their identification decision.
133 233 333	(WORST) These cases represent either a unanimous or majority ruling of observation code 3 from Table 9 which indicates an incorrect identification or non-identification decision by the examiner given ground truth data about the pair of images in question.

Following the identification decision step of the process, each examiner was asked to evaluate the image pair on the matter of fidelity loss. To aid in analysis and quantification of fidelity loss, the examiner's evaluation was collected by utilizing a Likert-type response scale [LIKERT]. The choices that the examiners were allowed to make are provided in Table 9 below. Observation codes are ordered in ascending order indicating a progressively greater amount of degradation from 1 to 4. Furthermore, the features summarized in observation code 3 are among those typically used for forensic-level decisions while features summarized in observation code 4 are those used primarily by automated-matchers in rendering a match decision.

Table 9 - Observation Codes for Image Degradation of Downsampling Treatments

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

Examiner responses were recorded by custom test apparatus consisting of a commodity computer and software designed and developed specifically for this study. More details on the equipment used are available in Appendix B. The examiners were not provided any time limits on their responses.

The responses from each of the 3 examiners were concatenated so that a response of 1, 3 and 2 becomes 132. This concatenated response was then sorted to list the values in increasing order, so 132 becomes 123. This process yields a response code that establishes a natural cost function as shown below in Table 10.

Table 10 - Sorted Examiner Response Code Classification

Responses from Observers	Classification
111 112 113 114	(BEST) These cases represent either a unanimous or majority ruling of observation code 1 from Table 9 indicating no apparent image degradation.
122 123 124 222 223 224	These cases represent either a unanimous or majority ruling of observation code 2 from Table 9. This case also includes the special split-decision cases of 1, 2, 3 and 1, 2, 4. These cases are considered borderline cases but are assigned to this bin as they are biased towards an acceptable image rating by two examiners noting little or no observable loss.
134 234 133 233 333 334	These cases represent either a unanimous or majority ruling of observation code 3 from Table 9. This case also includes the special split-decision cases of 1, 3, 4 and 2, 3, 4. These cases are considered biased-towards, and indicative of level-3 detail loss.
144 244 344 444	(WORST) These cases represent either a unanimous or majority ruling of observation code 4 from Table 9 indicating loss of both level-3 and level-2 feature detail.

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

The image pairs were presented in split-screen to the examiners in randomized order and randomized placement (left/right split screen placement) to mitigate potential order effects or positional bias (see Figure 9). Scientists overseeing the tests were blind to the placement order of the images, as well as to the downsampling methods represented in the image pairs. These factors were tracked by the test apparatus.

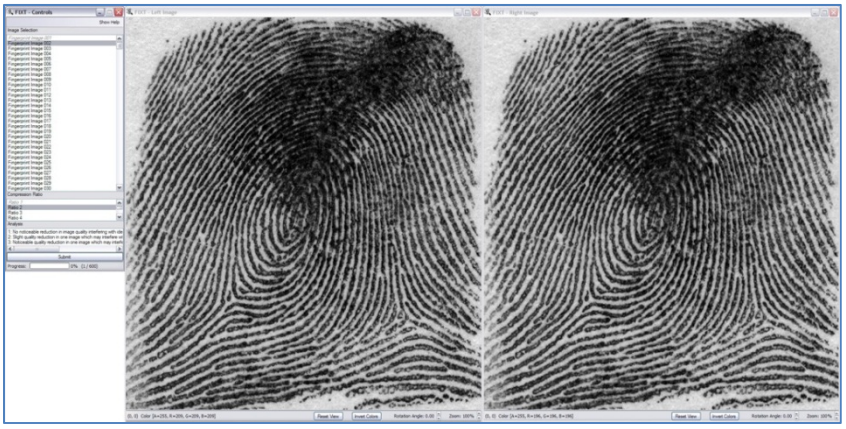


Figure 9 - Split-Screen Presentation of Image Pairs

Each of the 20,000 image pairs was guaranteed to be observed by three different examiners over the course of the study without repeating. Due to the physical limitations of even the best modern monitors, it is impossible to show a 500 ppi image without zooming or interpolation. The software apparatus enabled the examiner to view images at approximately 5x to 25x of the original size (see Appendix B for more information).

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

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

3.2.2. Automated Matching

The goal of the automated (machine) matching portion of the study was to determine how the downsampling treatments impacted an automated machine matcher's ability to match that image to a native-scanned 500 ppi image. This was done by generating minutiae templates for all images (native-scanned and downsampled images) and then generating match scores for the genuine match pairs. The following table is an example of a genuine match-pair:

Image source template	Subject ID	Finger position	Image type
probefile.xyt	01	03	500 ppi native scanned, WSQ compressed
galleryfile.xyt	01	03	1000 ppi scanned, downsampled to 500 ppi, WSQ compressed

For the machine matching process, two research applications from the NIST NBIS (NIST Biometric Image Software) distribution v 4.1.0 were utilized:

- mindtct.exe - this application generates a minutiae template according to the ANSI INCITS 378-2004 standard from a source image in either WSQ or PNG image format
- bozorth3.exe - this application generates a match score from an input consisting of two minutiae templates to be matched

The image sets used for all minutiae detection and subsequent, automated matching included:

- 500 ppi "native" images put through a single-cycle of WSQ compression using the NBIS distribution cwsq.exe application (v 4.1.0)
- Sample images that have been segmented from 10-print cards scanned at 1000 ppi, downsampled to 500 ppi using the treatments outlined in this document, and put through a single-cycle of WSQ compression using the NBIS distribution cwsq.exe application (v 4.1.0)

All minutiae templates were generated according to ANSI INCITS 378-2004 standard. The command line execution of mindtct is as follows:

- `mindtct -m1 fp-image-input template-output-root-filename`
  where -m1 specifies the output template to be ANSI INCITS 378-2004 standard compliant
  fp-image-input is the image of which to generate its minutiae template
  template-output-root-filename is the root filename of the generated minutiae template (file)

The reason that minutiae templates are generated for images is that the matcher application matches an image template and not the image itself.

All match scores for fingerprint pairs were processed and generated using the NIST Bozorth3 research matcher. The command-line execution of bozorth3 is as follows:

- bozorth3 -A outfmt=pgs -m1 probefile.xyt galleryfile.xyt
  where -A outfmt=pgs specifies that output lines will contain (s)core, (p)robe and (g)allery filename
  -m1 specifies the output template to be ANSI INCITS 378-2004 standard compliant
  probefile.xyt is the probe template filename
  galleryfile.xyt is the gallery template filename

For each treatment, 2039 match pairs were processed. The match scores were generated and captured for statistical analysis.

3.2.3. Spectral Image Validation/Verification Metric (SIVV)

Developed initially as a method to screen fingerprint databases for non-fingerprint images, segmentation errors, or mislabeled sample rates, the Spectral Image Validation Verification (SIVV) metric [LIBERT] provides a comparatively straightforward method by which to assess the frequency structure of an image. Pairwise display of the SIVV signals of images having undergone different processes enables summary visualization of the effects of the different processes across the composition frequency spectrum of the image. As a 1 dimensional representation of a 2 dimensional Fourier spectrum, the SIVV metric applied to a fingerprint image exhibits a peak corresponding to the frequency of the ridge spacing. Also, as shown in Figure 10, comparison of SIVV signals of a 500 ppi image and a corresponding downsampled 1000 ppi image shows the loss or gain of power over various frequencies.

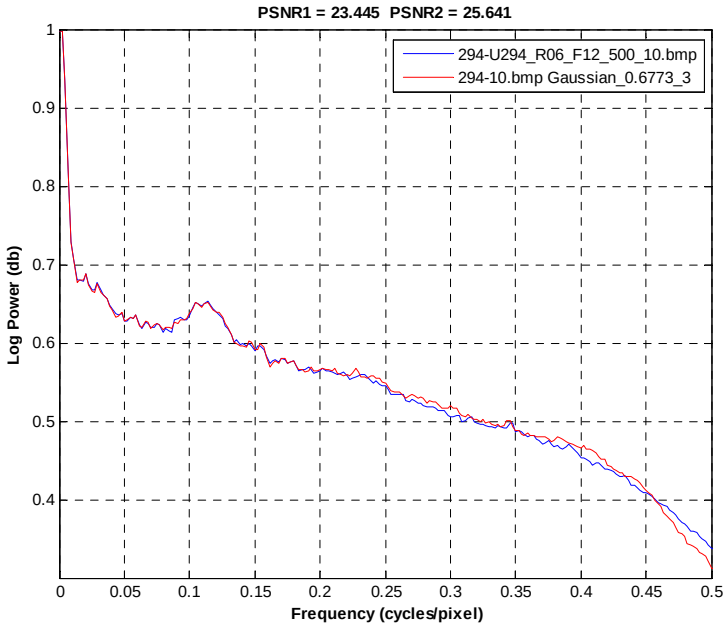


Figure 10 - NIST SIVV metric applied to a 500 ppi image having undergone 1-cycle of WSQ encoding/decoding and a corresponding 1000 ppi fingerprint encoded/decoded using JPEG2000, downsampled to 500 ppi, and encoded and decoded with WSQ.

The SIVV signals denoted as s_1 and s_2 are respectively vectors of SIVV signal values for images having undergone different processing methods under study. The frequency samples, *f*, in units of cycles per pixel correspond to image pixels or Fourier transform frequencies along the length of one half of the minimum dimension of the 2D Fourier transform of the image under examination. Frequency along this dimension is scaled to the interval [0, 0.5] cycles/pixel. Note that the power value at f=0 is the "direct current" (DC) term, corresponding to the average intensity of the image and is used to normalize the power spectrum.

Either differences or ratios of SIVV signals can provide quantitative measures for the comparison of compression methods. For the present study, we examine the Root Mean Squared difference (RMSD)[7] between the two SIVV signals, s_1 and s_2, over the entire frequency range 0 - 0.5 cycles/pixel.

$$RMSD(\mathbf{s}_1,\mathbf{s}_2)=\sqrt{\frac{\sum_{i=1}^{n}(\mathbf{s}_{1,i}-\mathbf{s}_{2,i})^2}{n}} \qquad (4)$$

where $n=|\mathbf{s}_1|=|\mathbf{s}_2|$ (i.e., the lengths of the signal vectors).

[7] It should be noted that RMSE and related image comparison measures such as Peak Signal to Noise Ratio (PSNR) were omitted from the analysis due to frequent differences in dimensions of downsampled 1000ppi images and 500ppi exemplars. Spectral, i.e. SIVV, comparison avoids difficulties of misregistration and size disparities. Hence, all mention of RMSD and correlation in this report pertains to comparison of SIVV signals.

The RMSD metric defined above can provide a measure of the overall difference between the SIVV spectra of downsampled 1000 ppi image and that of a corresponding image of the same fingerprint impression scanned at 500 ppi. In addition to global effects, the RMSD may be evaluated over smaller frequency intervals enabling the comparison of effects over frequency bands that may have particular relevance to fingerprint image quality or matching, as well as quantifying and isolating changes confined to bands that specifically impact either the machine matcher or expert examiners.

The RMSD measures the total deviation of point-wise comparison of the SIVV signals. The Pearson product moment correlation coefficient measures the parallelism between the two signals irrespective of the magnitude of the difference between them. Accordingly, we compute the correlation coefficient between $\mathbf{s}_1$ and $\mathbf{s}_2$ as

$$r(\mathbf{s}_1,\mathbf{s}_2)=\frac{\sum_{i=1}^{n}\left(\mathbf{s}_1-\bar{\mathbf{s}}_1\right)\left(\mathbf{s}_2-\bar{\mathbf{s}}_2\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathbf{s}_1-\bar{\mathbf{s}}_1\right)^2}\sqrt{\sum_{i=}^{n}\left(\mathbf{s}_2-\bar{\mathbf{s}}_2\right)^2}} \qquad \textbf{(5)}$$

where $\bar{\mathbf{s}}_1$ and $\bar{\mathbf{s}}_2$ are the arithmetic means of the two SIVV signal vectors.

4. Analysis

4.1. 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 (Frequency Spectrum Correlation of the Odd-Odd Decimation Case) in Figure 11 and Figure 12 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 utilized non-parametric statistical methods in the analysis of data collected in this study.

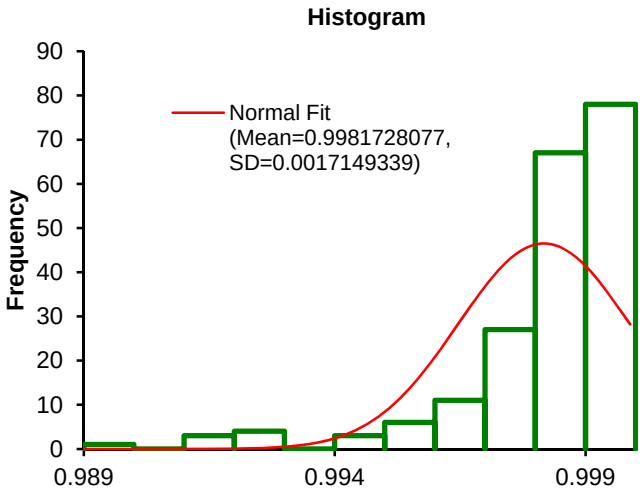


Figure 11 - Normal Fit Histogram, Frequency Spectrum Correlation, Decimation (Odd-Odd) Case

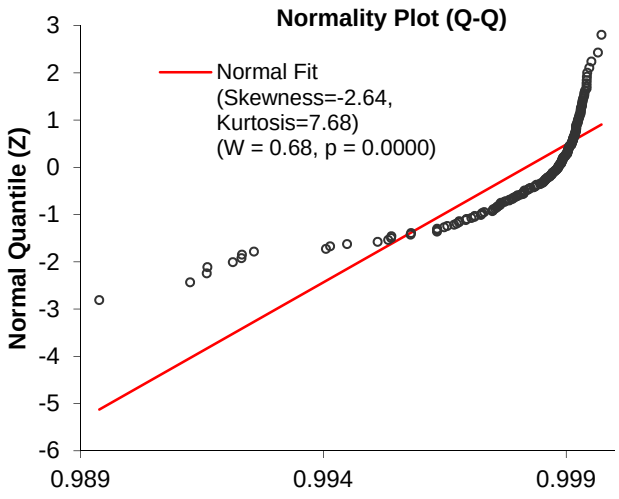


Figure 12 - Q-Q Normality Plot, Frequency Spectrum Correlation, Decimation (Odd-Odd) Case

4.2. Analysis Algorithm

As described in section 2, this study examines 28 treatments and 1 Control Case. Out of the 28 cases, 23 cases require decimation to perform geometric correction after processing, and one of the 28 cases is simply decimation without any further processing (see Section 5.7). Anecdotal evidence suggests that the decimation method should not result in any measurable change in the treatment pathway given the random distribution of the underlying data. Empirical evidence during an exploratory phase of this study demonstrated this to not be true, and that examination of decimation strategy alone showed measurable bias towards a particular decimation method. This was the primary reason why every possible combination of the four decimation strategy was examined for each of the treatments. At the time of the experiment the cause of this bias is not known. This bias is further explored in the Discussion and Future Work (section 7) of this study.

The pitfall in examining every combination of decimation strategy is that this combinatorial analysis amplifies out 24 of the 28 cases by a factor of 4, yielding 96 cases. The 96 cases involving decimation, combined with the 5 cases that did not require a decimation strategy yield a total of 101 cases that can overly complicate the analysis of the underlying problem.

Rather than comparing the full combination of treatments and sub-treatments together at the same time, a more simplified approach was devised to conduct the analysis in multiple stages.

In stage 1 of the analysis (See Figure 13) the 24 cases that each required 4 decimation sub-treatments (yielding 96 cases) had their sub-treatments ranked using Friedman's analysis. This allowed for the identification of the best and worst sub-treatments from the four possible sub-treatments.

In stage 2 of the analysis, the best decimation cases (sub-treatments) were down-selected and ranked using Friedman's analysis and in stage 3 of the analysis the worst decimation cases (sub-treatments) were down-selected and ranked using Friedman's analysis.

Determination of the best and worst cases allowed for the identification of treatment performance under the best decimation scenarios, as well as the worst decimation scenarios effectively creating a decision manifold that encompasses what is estimated to be the operational reality of the system.

While analysis of data primarily relies on mean ranks and Friedman's test, post-hoc analysis was performed at each stage of the analysis.

Post-hoc analysis [WILCOXON] was conducted in stage 1 of the analysis, but primarily for demonstration purposes and the results did not change selections based on mean rank. The results of this post-hoc analysis can however be used to identify other candidates that are "as good as" the rank 1 candidate (where there are no statistically significant differences in the post-hoc), in the event that an alternative rank 1 candidate is needed in practice for any reason. Post-hoc was also performed in stage 2 and stage 3 of the analysis to identify clustering of similar treatments. Additional information on this clustering algorithm has been provided in the results section (section 5) of this study.

For cases that required no decimation (and therefore had no sub-treatments), the singular case was propagated to phase two without any change.

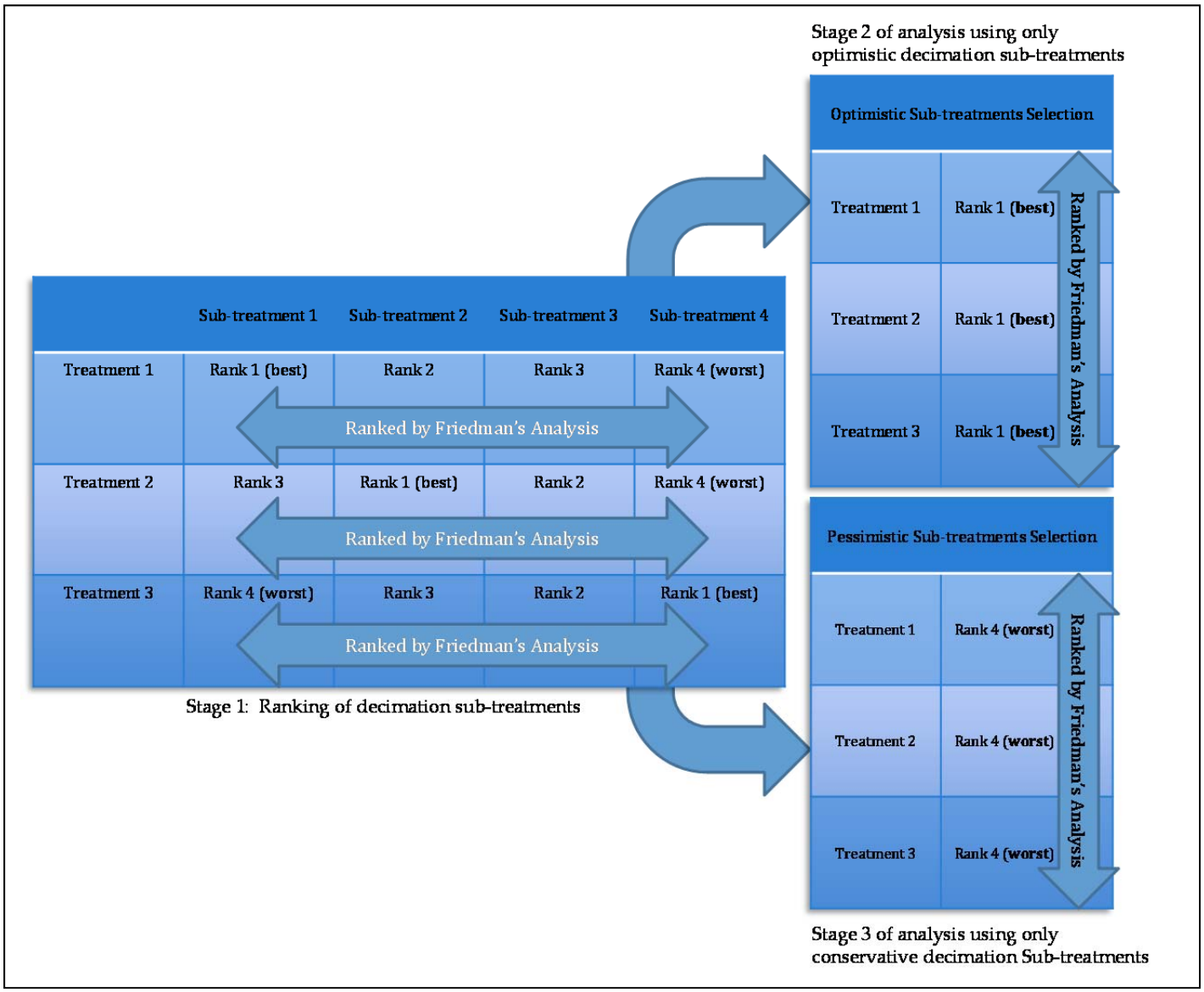


Figure 13 – Multi-Stage Analysis Overview

4.3. Hypothesis Testing

4.3.1. Friedman's Procedure

The Friedman test [CONOVER] is a non-parametric test for analyzing randomized complete block designs and is an extension of the sign test when there may be more than two treatments of interest. The Friedman test assumes that there are two or more experimental treatments (k). The observations are arranged in b blocks as in the example shown in Table 11.

Table 11 - Block Design Example

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

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

$$R_j = \sum_{i=1}^{b} R\left(X_{i,}\right) \tag{6}$$

for $j = 1, 2, \ldots, k$.

Then the Friedman test is

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

where $\mu_{j\,is}$ the mean rank of treatment j, j = 1 … k.

Test Statistic:

$$T_1 = \frac{12}{bk(k+1)} \sum_{i=1}^{k} \left(\left(R_i - b(k+1) \right) / 2 \right)^2 \tag{7}$$

If there are ties, then

$$T_1 = \frac{(k-1) \sum_{i=1}^{k} \left(R_i - \frac{b(k+1)}{2} \right)^2}{A_1 - C_1} \tag{8}$$

where

$$A_1 = \sum_{i=1}^{b} \sum_{j=1}^{k} \left(R\left(X_{ij}\right) \right)^2 \tag{9}$$

$$C_1 = \frac{bk(k+1)^2}{4} \tag{10}$$

Note that Conover recommends the statistic

$$T_2 = \frac{(b-1)T_1}{b(k-1) - T_1} \tag{11}$$

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

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

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

$$T_2 > F_{(\alpha,k-1,(b-1)(k-1))} \quad \textbf{(12)}$$

where F is the percent point function of the F distribution.

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

$$T_1 = \mathrm{X}^2_{(\alpha,k-1)} \quad \textbf{(13)}$$

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

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

4.3.2. Wilcoxon Signed Rank Test

Hypothesis testing was accomplished in two phases. For each assemblage of downsampling methods, the Friedman test was performed to a) determine if any of the methods differed from at least one of the other methods; and b) to determine the set of mean rank scores of the downsampling methods under consideration and to sort the methods from "best performing" to "worst performing" on the basis of the mean rank scores. Commonly, Friedman's test is followed by *post hoc* analysis involving multiple pairwise comparisons to determine which of the treatments differ from one another at levels of statistical significance. For purposes of the present study, however, the multiple pairwise comparison procedure was used only to identify which of the cluster of alternative methods near the top of the mean rank score list differed significantly from that method occupying the top position.

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

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

Given the ordering of the downsampling methods as indicated by the mean rank scores, the Wilcoxon test was applied to the columns of the original data matrix so as to successively compare pairwise measurements of the treatment deemed the best with those deemed next best, and so on down the list until a significant difference was found. Thus, we were able to assess if a method deemed best by the mean rank criterion was statistically significantly different from potential contenders. Moreover, the ordering of the methods with successive pairwise testing obviated the need to test every method against every other method. That limiting the number of comparisons is important will become apparent from the discussion which follows.

Use of *post hoc* analysis requires the use of some measure to control the experiment-wise Type I error rate, i.e., the probability of incorrectly rejecting the null hypothesis. The philosophical underpinnings are explained in [CURRAN-EVERETT], but essentially the idea is that where an experiment consists of multiple decisions, each carrying a probability of error, the likelihood of making an incorrect decision increases with the number of decisions. Accordingly, for an experiment involving multiple planned, dependent (non-orthogonal) decisions to accept or reject the null hypothesis, the accepted error probability for incorrectly rejecting the null hypothesis must be apportioned among the various decisions such that the overall error probability does not exceed the acceptable error rate. Various methods for controlling this "family-wise" Type I error are described in [CONOVER] and [SOKAL].

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

$$\alpha'_k = \frac{\alpha}{k} \quad \textbf{(14)}$$

where *k* = the number of pairwise comparisons. Hence, given a set of downsampling treatments ordered by mean rank score, comparison of each of several contending methods may be compared to the method in the top position via the Wilcoxon test which tests the probability that the two methods yield equivalent results. For each comparison, we adjust the threshold for acceptance of the null hypothesis according to equation (14). We are not interested in all pairwise comparisons, but at most only in comparisons involving the top ranking method of each of other methods of a group. Thus, we proceed down the list of methods ordered by mean rank to test the difference between the top ranking

method with others until we find a statistically significant difference. We group these methods as statistically undifferentiable. The method that was found to be statistically different from the top ranked of the grouping just formed becomes the top ranked of the next group to which we compare with methods below it in the list of mean ranks. Again, we form a group of methods when we encounter a significant difference. We proceed similarly to the end of the list. At the end of the procedure we have a set of groups members of which are statistically similar to their top ranked member. We employ this rather elaborate procedure in order to facilitate integration of analyses of the various metrics used in the study as will be discussed later in section 5.

Some of the pairwise tests involved at most three comparisons with a top ranking method. For these pairwise tests we set the number of comparisons at $k = 3$ and use $\alpha' = 0.05/3 = 0.0167$ as the threshold for rejection of the null hypothesis in each pairwise test. For other pairwise tests, we know that there will be at most comparison of 28 methods with a top ranking method, but we do not know *a priori* the number of methods that might form a grouping of methods. Accordingly, we allow that we could have a maximum of 28 pairwise comparisons, i.e. $k = 28$ and conservatively use the adjusted alpha value $\alpha' = 0.05/28 = 0.0018$ as the threshold for rejection of the null hypothesis in each pairwise test.

5. Results

5.1. Investigative Goal 1: Expert Examiner Quality Assessment

5.1.1. Investigative Analysis 1

Data collection from the examiners yielded a series of 200 sorted response codes for each treatment and sub-treatment as described in Table 3. The data was coded and ranged from 111 (best) to 444 (worst) to represent the three examiner decisions as described in Table 10 and the data was arranged in a table. In the first phase of analysis, the Friedman's statistical test was performed on each treatment that required a decimation sub-treatment. This resulted in rankings of the sub-treatments (see Table 12). This allowed for the identification of the sub-treatment from best (lowest mean rank) to worst (highest mean rank), and provided us an indication that at least one statistically significant difference in distributions exists between the four sub-treatments via Friedman's *p*. The ranking of the sub-treatments allows us to select the best sub-treatment for each treatment and to propagate that to the next phase of analysis. In addition to selecting the top ranked sub-treatment per the rank analysis, a Wilcoxon post-hoc analysis was performed between treatments having Rank 1 and Rank 2, Rank 1 and Rank 3 and Rank 1 and Rank 4 to determine if the top ranked sub-treatment is indeed different from Rank 2/3/4.

Table 12 - Expert Examiner Quality Data Analysis, Stage One (Sub-treatment Pare-Down, Adjusted α = 0.0167)

Treatment	Sub-treatment, Friedman's Mean Rank, Lower is Better								
	Even-Even	Even-Odd	Odd-Even	Odd-Odd	Friedman's p	X^2 statistic	p, Wilcoxon Rank 1 ≠ 2	p, Wilcoxon Rank 1 ≠ 3	p, Wilcoxon Rank 1 ≠ 4
Native 500 ppi	-	-	-	-	-	-	-	-	
Averaging	-	-	-	-	-	-	-	-	
NIST Spectral Truncation	-	-	-	-	-	-	-	-	
Decimation	3.39	1.96	2.75	1.91	<0.0001	203.93	0.7153	<0.0001	
Transcoded A	-	-	-	-	-	-	-	-	
Transcoded B	-	-	-	-	-	-	-	-	
Gaussian σ = 0.3369, r = 2	3.09	1.81	3.14	1.97	<0.0001	200.78	0.1925	<0.0001	
Gaussian σ = 0.4220, r = 2	2.99	1.96	3.05	2.01	<0.0001	145.58	0.1066	<0.0001	
Gaussian σ = 0.5071, r = 3	2.80	2.07	3.15	1.98	<0.0001	133.05	0.4056	<0.0001	
Gaussian σ = 0.5642, r = 3	2.55	2.29	3.00	2.17	<0.0001	55.48	0.2105	0.0003	
Gaussian σ = 0.5922, r = 3	2.64	2.11	3.00	2.25	<0.0001	68.09	0.2635	<0.0001	
Gaussian σ = 0.6773, r = 3	2.36	2.30	2.98	2.36	<0.0001	42.06	0.7219	0.6156	<0.0001
Gaussian σ = 0.7624, r = 3	2.37	2.31	2.97	2.35	<0.0001	40.40	0.5424	0.5122	<0.0001
Gaussian σ = 0.8475, r = 4	2.30	2.60	2.85	2.25	<0.0001	32.66	0.8418	0.0092	
Gaussian σ = 0.9326, r = 4	2.29	2.66	2.81	2.24	<0.0001	33.35	0.8664	0.0008	
Gaussian σ = 1.0177, r = 4	2.23	2.81	2.70	2.26	<0.0001	37.51	0.8588	0.0013	
Gaussian σ = 1.1028, r = 4	2.25	2.80	2.65	2.30	<0.0001	29.62	0.5904	0.0091	
Gaussian σ = 1.1879, r = 5	2.23	2.76	2.78	2.24	<0.0001	40.31	0.8136	<0.0001	
Gaussian σ = 1.2730, r = 5	2.22	2.94	2.44	2.41	<0.0001	40.09	0.0920	0.4730	<0.0001
Gaussian σ = 1.3581, r = 5	2.24	2.93	2.31	2.52	<0.0001	42.72	0.1528	0.0127	
Gaussian σ = 1.4432, r = 5	2.33	2.91	2.24	2.52	<0.0001	39.61	0.0097		
Gaussian σ = 1.5283, r = 6	2.40	2.89	2.12	2.59	<0.0001	45.90	<0.0001		
Gaussian σ = 1.6134, r = 6	2.49	2.80	2.07	2.64	<0.0001	43.21	<0.0001		
Gaussian σ = 1.6985, r = 6	2.61	2.88	1.85	2.66	<0.0001	90.54	<0.0001		
Gaussian σ = 1.7836, r = 6	2.59	2.84	1.99	2.58	<0.0001	58.95	<0.0001		
Gaussian σ = 1.8687, r = 7	2.64	2.85	1.87	2.64	<0.0001	84.63	<0.0001		
Gaussian σ = 1.9538, r = 7	2.62	2.77	1.98	2.63	<0.0001	57.46	<0.0001		
Gaussian σ = 2.0389, r = 7	2.60	2.78	1.92	2.71	<0.0001	71.89	<0.0001		
Gaussian σ = 2.1240, r = 7	2.66	2.73	2.00	2.62	<0.0001	52.85	<0.0001		

In stage two of analysis, the Rank 1 candidates from each sub-treatment were compared against each other as well as against the cases that received no sub-treatment[8]. The data from this phase of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list. Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. In each comparison the higher-ranked of two of the ordinal cases was made the exemplar case, P0 (case no.) compared with a lower-ranked case, P1 (case no.). These comparisons of P0 to P1 were continued until a statistically significant difference was observed at which point the significantly different case was made the new P0 and comparisons continued with lower-ranked cases until the condition P0 ≠ P1 (in the sense of statistical significance) became true. The cases that did not demonstrate a statistically significant difference were clustered into one singular rank, the Clustered Rank Table of Table 13. This process continued until a list of

[8] i.e., the treatment did not include decimation as part of its processing.

clustered of rank values was established, with each clustered rank representing data cases that are not significantly different from the others (and hence clustered as such). Each clustered group was then provided an ordinal clustered rank value.

Table 13 - Expert Examiner Quality Data Analysis, Stage Two (Optimistic Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank Sum	Mean Rank	Case No.	P_0 (Case No.)	P_1 (Case No.)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Gaussian σ = 0.5922, r = 3, Even-Odd	1725.5	8.63	1	NA	NA	NA		1
Gaussian σ = 0.5071, r = 3, Odd-Odd	1820.0	9.10	2	1	2	0.9043		1
Gaussian σ = 0.8475, r = 4, Odd-Odd	1832.0	9.16	3	1	3	0.4182		1
Gaussian σ = 0.6773, r = 3, Even-Odd	1853.5	9.27	4	1	4	0.5092		1
Gaussian σ = 0.7624, r = 3, Even-Odd	1962.5	9.81	5	1	5	0.1837		1
Gaussian σ = 0.9326, r = 4, Odd-Odd	1969.5	9.85	6	1	6	0.0170		1
Gaussian σ = 1.0177, r = 4, Even-Even	1976.0	9.88	7	6	7	0.0262		1
Gaussian σ = 0.5642, r = 3, Odd-Odd	1986.0	9.93	8	6	8	0.2635		1
Gaussian σ = 0.3369, r = 2, Even-Odd	2010.0	10.05	9	6	9	0.1305		1
Gaussian σ = 0.4220, r = 2, Even-Odd	2063.0	10.32	10	6	10	0.0346		1
Native 500 ppi	2213.0	11.07	11	6	11	0.0011	Yes	2
Gaussian σ = 1.1028, r = 4, Even-Even	2294.5	11.47	12	6	12	0.3897		2
Spectral Truncation	2305.0	11.53	13	6	13	0.4197		2
Gaussian σ = 1.1879, r = 5, Even-Even	2731.5	13.66	14	6	14	0.0015	Yes	3
Decimation, Odd-Odd	2922.5	14.61	15	14	15	0.6203		3
Gaussian σ = 1.2730, r = 5, Even-Even	3032.5	15.16	16	14	16	0.0147		3
Transcoded B	3038.5	15.19	17	14	17	0.0340		3
Gaussian σ = 1.3581, r = 5, Even-Even	3389.0	16.95	18	14	18	<0.0001	Yes	4
Gaussian σ = 1.4432, r = 5, Odd-Even	3580.5	17.90	19	18	19	0.8652		4
Averaging	3628.0	18.14	20	18	20	0.8508		4
Gaussian σ = 1.5283, r = 6, Odd-Even	3655.0	18.28	21	18	21	0.5779		4
Gaussian σ = 1.6134, r = 6, Odd-Even	3957.0	19.79	22	18	22	0.1077		4
Transcoded A	4045.5	20.23	23	18	23	0.0047		4
Gaussian σ = 1.6985, r = 6, Odd-Even	4121.5	20.61	24	18	24	0.0094		4
Gaussian σ = 1.7836, r = 6, Odd-Even	4404.5	22.02	25	18	25	<0.0001	Yes	5
Gaussian σ = 1.8687, r = 7, Odd-Even	4426.0	22.13	26	25	26	0.9484		5
Gaussian σ = 1.9538, r = 7, Odd-Even	4558.5	22.79	27	25	27	0.0305		5
Gaussian σ = 2.0389, r = 7, Odd-Even	4711.5	23.56	28	25	28	0.0024		5
Gaussian σ = 2.1240, r = 7, Odd-Even	4787.5	23.94	29	25	29	<0.0001	Yes	6

n	200	
Friedman's statistic	2187.36	
DF	28	
Friedman's p	<0.0001	(X^2 approximation, corrected for ties)

In stage three of analysis, the Rank 4 (worst) decimation candidates from each sub-treatment were compared against each other as well as the cases that received no sub-treatment. The data from this stage of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 14). Post-hoc analysis was again performed on this list, comparing Case 1 to Case 2 and so on. As described above, this comparison continued until a statistically significant difference was observed. The values that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered rank values was established, with each clustered rank representing data cases that are not significantly different from the others and clustered as such. Each clustered group was again provided an ordinal clustered rank value.

Table 14 - Expert Examiner Quality Data Analysis, Stage Three (Conservative Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	P_0 (Case No.)	P_1 (Case No.)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Spectral Truncation	1513.5	7.57	1	NA	NA	NA		1
Native 500 ppi	1531.0	7.66	2	1	2	0.4197		1
Gaussian σ = 0.9326, r = 4, Odd-Even	1847.0	9.24	3	1	3	0.0621		1
Gaussian σ = 1.0177, r = 4, Even-Odd	1857.0	9.29	4	1	4	0.0355		1
Gaussian σ = 0.8475, r = 4, Odd-Even	1931.0	9.66	5	1	5	0.0896		1
Gaussian σ = 0.6773, r = 3, Odd-Even	1993.5	9.97	6	1	6	0.0147		1
Gaussian σ = 0.7624, r = 3, Odd-Even	2051.0	10.26	7	1	7	0.0011	Yes	2
Transcoded B	2060.0	10.30	8	7	8	0.0146		2
Gaussian σ = 1.1028, r = 4, Even-Odd	2085.5	10.43	9	7	9	0.1849		2
Gaussian σ = 0.5922, r = 3, Odd-Even	2182.0	10.91	10	7	10	0.2023		2
Gaussian σ = 0.5642, r = 3, Odd-Even	2304.5	11.52	11	7	11	0.1107		2
Gaussian σ = 1.1879, r = 5, Odd-Even	2346.0	11.73	12	7	12	0.0209		2
Gaussian σ = 0.5071, r = 3, Odd-Even	2613.0	13.07	13	7	13	0.0006	Yes	3
Gaussian σ = 0.4220, r = 2, Odd-Even	2785.5	13.93	14	13	14	0.4410		3
Averaging	2832.5	14.16	15	13	15	0.2342		3
Gaussian σ = 1.2730, r = 5, Even-Odd	3016.0	15.08	16	13	16	<0.0001	Yes	4
Gaussian σ = 0.3369, r = 2, Odd-Even	3139.5	15.70	17	16	17	0.0598		4
Gaussian σ = 1.3581, r = 5, Even-Odd	3243.5	16.22	18	16	18	0.0342		4
Transcoded A	3374.5	16.87	19	16	19	0.8303		4
Gaussian σ = 1.4432, r = 5, Even-Odd	3523.0	17.62	20	16	20	<0.0001	Yes	5
Gaussian σ = 1.5283, r = 6, Even-Odd	3755.0	18.78	21	20	21	0.0097		5
Gaussian σ = 1.6134, r = 6, Even-Odd	3933.0	19.67	22	20	22	<0.0001	Yes	6
Decimation Even-Even	3962.5	19.81	23	22	23	0.0052		6
Gaussian σ = 1.6985, r = 6, Even-Odd	4216.0	21.08	24	22	24	0.0046		6
Gaussian σ = 1.7836, r = 6, Even-Odd	4329.5	21.65	25	22	25	0.0003	Yes	7
Gaussian σ = 1.8687, r = 7, Even-Odd	4559.5	22.80	26	25	26	0.0072		7
Gaussian σ = 1.9538, r = 7, Even-Odd	4593.0	22.97	27	25	27	0.0009	Yes	8
Gaussian σ = 2.0389, r = 7, Even-Odd	4661.5	23.31	28	27	28	0.2481		8
Gaussian σ = 2.1240, r = 7, Even-Odd	4760.5	23.80	29	27	29	0.0309		8

n	200	
Friedman's statistic	2210.35	
DF	28	
Friedman's p	<0.0001	(X^2 approximation, corrected for ties)

5.1.2. Investigative Result 1

Based on Friedman's analysis of the pared down cases, the Gaussian filter with the parameters of σ = 0.5922, radius = 3 utilizing the Even-Odd decimation strategy has the best mean rank among the methods utilizing optimistic decimation strategy (stage 2 analysis). Post-hoc analysis using Wilcoxon indicates that this method is equally as effective and has a score distribution that is not significantly different from methods (Gaussian σ = 0.5071, r = 3, Odd-Odd), (Gaussian σ = 0.8475, r = 4, Odd-Odd), (Gaussian σ = 0.6773, r = 3, Even-Odd) and (Gaussian σ = 0.7624, r = 3, Even-Odd) therefore these five methods share the top rank and are clustered as the top ranking methods as shown in Table 13.

Analysis using a conservative decimation strategy (stage 3 analysis) showed that Spectral Truncation has the best mean rank. Post-hoc analysis using Wilcoxon indicates that this method is equally as effective and has a score distribution that is not significantly different from methods (Gaussian σ = 0.9326, r = 4, Odd-Even), (Gaussian σ = 1.0177, r = 4, Even-Odd), (Gaussian σ = 0.8475, r = 4, Odd-Even) and (Gaussian σ = 0.6773, r = 3, Odd-Even). Post-hoc analysis also showed that all of these methods have a score distribution that is not significantly different from the Control Case of native 500 ppi images.

5.2. Investigative Goal 2: Expert Examiner Identification Assessment

5.2.1. Investigative Analysis 2

Data collection from the examiners yielded a series of 200 sorted response codes for each treatment and sub-treatment as described in Table 3. The data was coded and ranged from 111 (best) to 333 (worst) to represent the three examiner identification decisions. Statistical analysis showed that there was no significant distinction between the decimation sub-treatments of the 24 cases that required decimation as shown in Table 15.

Table 15 - Expert Examiner Match Data Analysis, Stage One (Sub-treatment Pare-Down, Adjusted α = 0.0167)

Treatment	Sub-treatment, Friedman's Mean Rank, Lower is Better								
	Even-Even	Even-Odd	Odd-Even	Odd-Odd	Friedman's p	X^2 statistic	p, Wilcoxon Rank 1 ≠ 2	p, Wilcoxon Rank 1 ≠ 3	p, Wilcoxon Rank 1 ≠ 4
Native 500 ppi	-	-	-	-	-	-	-	-	-
Averaging	-	-	-	-	-	-	-	-	-
NIST Spectral Truncation	-	-	-	-	-	-	-	-	-
Decimation	2.51	2.50	2.50	2.50	0.3916	3.00	-	-	-
Transcoded A	-	-	-	-	-	-	-	-	-
Transcoded B	-	-	-	-	-	-	-	-	-
Gaussian σ = 0.3369, r = 2	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 0.4220, r = 2	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 0.5071, r = 3	2.50	2.51	2.51	2.50	0.5724	2.00	-	-	-
Gaussian σ = 0.5642, r = 3	2.50	2.50	2.50	2.51	0.3916	3.00	-	-	-
Gaussian σ = 0.5922, r = 3	2.51	2.50	2.50	2.51	0.5724	2.00	-	-	-
Gaussian σ = 0.6773, r = 3	2.50	2.50	2.51	2.50	0.3916	3.00	-	-	-
Gaussian σ = 0.7624, r = 3	2.50	2.51	2.50	2.50	0.3916	3.00	-	-	-
Gaussian σ = 0.8475, r = 4	2.50	2.50	2.50	2.51	0.3916	3.00	-	-	-
Gaussian σ = 0.9326, r = 4	2.50	2.50	2.50	2.50	1.0000	3.00	-	-	-
Gaussian σ = 1.0177, r = 4	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 1.1028, r = 4	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 1.1879, r = 5	2.50	2.51	2.51	2.50	0.5724	2.00	-	-	-
Gaussian σ = 1.2730, r = 5	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 1.3581, r = 5	2.50	2.51	2.50	2.50	0.3916	3.00	-	-	-
Gaussian σ = 1.4432, r = 5	2.50	2.51	2.50	2.51	0.5724	2.00	-	-	-
Gaussian σ = 1.5283, r = 6	2.49	2.49	2.50	2.51	0.2998	3.67	-	-	-
Gaussian σ = 1.6134, r = 6	2.49	2.51	2.49	2.50	0.2998	3.67	-	-	-
Gaussian σ = 1.6985, r = 6	2.50	2.50	2.50	2.50	1.0000	0.00	-	-	-
Gaussian σ = 1.7836, r = 6	2.50	2.52	2.50	2.50	0.1116	6.00	-	-	-
Gaussian σ = 1.8687, r = 7	2.51	2.52	2.49	2.49	0.1447	5.40	-	-	-
Gaussian σ = 1.9538, r = 7	2.50	2.50	2.51	2.51	0.5724	2.00	-	-	-
Gaussian σ = 2.0389, r = 7	2.50	2.50	2.51	2.50	0.3916	3.00	-	-	-
Gaussian σ = 2.1240, r = 7	2.51	2.49	2.50	2.50	0.5724	2.00	-	-	-

In the next stage of analysis, the Rank 1 candidates from each sub-treatment were compared against each other and against the cases that received no sub-treatment. Since the data from phase one of this analysis indicated no statistically significant difference between the sub-treatments, the sub-treatment with the lowest mean ranks were selected for phase two of the analysis. Where a tie existed in the mean rank, an arbitrary selection was made. The rationale for this is that in cases where a statistically significant difference cannot be observed, nor an anecdotal difference can be observed, then the cases can be treated as operationally equivalent for the purposes of phase one of this analysis.

The data from this phase of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal rank ranging from 1 to 29 was assigned to this sorted list (see Table 16). As with the previous stage of analysis, Friedman's test of the values here indicates no significant differences (p=0.2060). Since no significant differences were shown in stage one of analysis between the different decimation cases, only stage 2 of the analysis was performed and stage 3 was skipped as there were no demonstrated differences between the optimistic and conservative decimation strategies.

Table 16 - Expert Examiner Match Data Analysis, Stage Two (Optimistic Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	μ_0 (Case No.)	μ_1 (Case No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
Averaging	2993.5	14.97	1	-	-	-	-	-
Decimation, Even-Odd	2993.5	14.97	2	-	-	-	-	-
Gaussian σ = 0.3369, r = 2, Even-Even	2993.5	14.97	3	-	-	-	-	-
Gaussian σ = 0.4220, r = 2, Even-Odd	2993.5	14.97	4	-	-	-	-	-
Gaussian σ = 0.5071, r = 3, Odd-Odd	2993.5	14.97	5	-	-	-	-	-
Gaussian σ = 0.5922, r = 3, Even-Odd	2993.5	14.97	6	-	-	-	-	-
Gaussian σ = 0.6773, r = 3, Odd-Odd	2993.5	14.97	7	-	-	-	-	-
Gaussian σ = 0.9326, r = 4, Odd-Odd	2993.5	14.97	8	-	-	-	-	-
Gaussian σ = 1.0177, r = 4, Even-Even	2993.5	14.97	9	-	-	-	-	-
Gaussian σ = 1.1028, r = 4, Even-Odd	2993.5	14.97	10	-	-	-	-	-
Gaussian σ = 1.1879, r = 5, Odd-Odd	2993.5	14.97	11	-	-	-	-	-
Gaussian σ = 1.2730, r = 5, Odd-Even	2993.5	14.97	12	-	-	-	-	-
Gaussian σ = 1.3581, r = 5, Odd-Odd	2993.5	14.97	13	-	-	-	-	-
Gaussian σ = 1.4432, r = 5, Odd-Even	2993.5	14.97	14	-	-	-	-	-
Gaussian σ = 1.6134, r = 6, Odd-Even	2993.5	14.97	15	-	-	-	-	-
Gaussian σ = 1.6985, r = 6, Odd-Odd	2993.5	14.97	16	-	-	-	-	-
Gaussian σ = 1.7836, r = 6, Even-Even	2993.5	14.97	17	-	-	-	-	-
Gaussian σ = 1.8687, r = 7, Odd-Even	2993.5	14.97	18	-	-	-	-	-
Transcoded A	2993.5	14.97	19	-	-	-	-	-
Gaussian σ = 0.5642, r = 3, Even-Odd	3008.0	15.04	20	-	-	-	-	-
Gaussian σ = 0.7624, r = 3, Even-Even	3008.0	15.04	21	-	-	-	-	-
Gaussian σ = 0.8475, r = 4, Even-Odd	3008.0	15.04	22	-	-	-	-	-
Gaussian σ = 1.5283, r = 6, Even-Even	3008.0	15.04	23	-	-	-	-	-
Gaussian σ = 1.9538, r = 7, Even-Odd	3008.0	15.04	24	-	-	-	-	-
Gaussian σ = 2.0389, r = 7, Odd-Odd	3008.0	15.04	25	-	-	-	-	-
Gaussian σ = 2.1240, r = 7, Even-Odd	3008.0	15.04	26	-	-	-	-	-
Transcoded B	3008.0	15.04	27	-	-	-	-	-
Spectral Truncation	3022.5	15.11	28	-	-	-	-	-
Native 500 ppi	3037.0	15.19	29	-	-	-	-	-
Friedman's statistic	33.85							
DF	28							
p	0.2060	(X2 approximation, corrected for ties)						

5.2.2. Investigative Result 2

Based on the analysis of the data, it is evident that the downsampling methods examined did not appear to cause a significant measurable impact on examiner objective analysis, and therefore their ability to make an identification decision was not impacted even though they could detect degradation as a result of the various downsampling methods as shown in the subjective analysis section of this study.

5.3. Investigative Goal 3: Automated Matching Assessment

5.3.1. Investigative Analysis 3

Data collection by matching the processed image to the unprocessed native image using the NIST research matcher (Bozorth3) can yield match scores ranging from 0 to 65535, with higher scores indicative of a higher probability of a true match as determined by the machine. In the first phase of analysis, Friedman's statistical test was performed on each treatment that required the decimation sub-treatment. This resulted in Friedman's ranking of the sub-treatments (see Table 17). This allowed for the stratification of the sub-treatment from best (highest mean rank) to worst (lowest mean rank), and provided us an indication that at least one statistically significant difference in distributions exists between the four sub-treatments via Friedman's *p*. The ranking of the sub-treatments allows us to select the best sub-treatment for each treatment and propagate that to the next phase of analysis. In addition to selecting the top ranked sub-treatment per the rank analysis, a Wilcoxon post-hoc analysis was performed between match data of methods having Rank 1 and Rank 2, Rank 1 and Rank 3 and Rank 1 and Rank 4 to determine if the top ranked sub-treatment is indeed different from that of Ranks 2, 3, and 4.

Table 17 - Automated Matching Data Analysis, Stage One (Sub-treatment Pare-Down, Adjusted α = 0.0167)

Treatment	Sub-treatment, Friedman's Mean Rank, Lower is Better								
	Even-Even	Even-Odd	Odd-Even	Odd-Odd	Friedman's p	X^2 statistic	p, Wilcoxon Rank 1 ≠ 2	p, Wilcoxon Rank 1 ≠ 3	p, Wilcoxon Rank 1 ≠ 4
Native 500 ppi	-	-	-	-	-	-	-		-
Averaging	-	-	-	-	-	-	-		-
Spectral Truncation	-	-	-	-	-	-	-	-	-
Decimation	2.14	2.70	2.51	2.65	<0.0001	238.45	0.0869	<0.0001	-
Transcoded A	-	-	-	-	-	-	-	-	-
Transcoded B	-	-	-	-	-	-	-	-	-
Gaussian σ = 0.3369, r = 2	2.14	2.72	2.45	2.69	<0.0001	264.68	0.0812	<0.0001	-
Gaussian σ = 0.4220, r = 2	2.21	2.70	2.50	2.58	<0.0001	159.74	0.0026	-	-
Gaussian σ = 0.5071, r = 3	2.39	2.59	2.38	2.64	<0.0001	66.24	<0.0001	-	-
Gaussian σ = 0.5642, r = 3	2.42	2.38	2.66	2.55	<0.0001	60.96	0.0061	-	-
Gaussian σ = 0.5922, r = 3	2.47	2.37	2.62	2.53	<0.0001	40.94	0.0201	<0.0001	-
Gaussian σ = 0.6773, r = 3	2.40	2.58	2.61	2.41	<0.0001	46.77	0.5062	<0.0001	-
Gaussian σ = 0.7624, r = 3	2.48	2.53	2.62	2.37	<0.0001	38.33	0.0178	<0.0001	-
Gaussian σ = 0.8475, r = 4	2.49	2.50	2.66	2.35	<0.0001	62.57	<0.0001	-	-
Gaussian σ = 0.9326, r = 4	2.54	2.47	2.61	2.38	<0.0001	35.86	0.0645	<0.0001	-
Gaussian σ = 1.0177, r = 4	2.58	2.46	2.61	2.35	<0.0001	54.57	0.6809	0.0001	-
Gaussian σ = 1.1028, r = 4	2.59	2.43	2.63	2.35	<0.0001	65.23	0.3539	<0.0001	-
Gaussian σ = 1.1879, r = 5	2.60	2.34	2.65	2.42	<0.0001	80.16	0.2758	<0.0001	-
Gaussian σ = 1.2730, r = 5	2.59	2.31	2.73	2.37	<0.0001	137.55	0.0015	-	-
Gaussian σ = 1.3581, r = 5	2.59	2.33	2.71	2.37	<0.0001	120.25	0.0059	-	-
Gaussian σ = 1.4432, r = 5	2.59	2.28	2.73	2.40	<0.0001	152.61	0.0015	-	-
Gaussian σ = 1.5283, r = 6	2.61	2.26	2.74	2.39	<0.0001	172.40	0.0009	-	-
Gaussian σ = 1.6134, r = 6	2.61	2.25	2.70	2.44	<0.0001	144.72	<0.0001	-	-
Gaussian σ = 1.6985, r = 6	2.63	2.26	2.69	2.42	<0.0001	143.08	0.0457	<0.0001	-
Gaussian σ = 1.7836, r = 6	2.67	2.20	2.43	2.70	<0.0001	202.28	<0.0001	-	-
Gaussian σ = 1.8687, r = 7	2.61	2.22	2.74	2.43	<0.0001	189.73	0.0018	<0.0001	-
Gaussian σ = 1.9538, r = 7	2.64	2.19	2.72	2.46	<0.0001	202.75	0.0030	-	-
Gaussian σ = 2.0389, r = 7	2.63	2.20	2.71	2.46	<0.0001	194.07	0.0078	-	-
Gaussian σ = 2.1240, r = 7	2.67	2.16	2.72	2.45	<0.0001	245.84	0.0572	<0.0001	-

In the next stage of analysis, the Rank 1 candidates from each sub-treatment were compared against each other and the cases that received no sub-treatment. The data from this phase of analysis was sorted according to mean rank, from largest (best) to smallest and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 18). Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. As described in 5.1, this comparison continued until a statistically significant difference was observed. The values that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered rank of values was established, with each clustered rank representing data cases that are not significantly different from the others (and hence are clustered as such).

Table 18 - Automated Matching Data Analysis, Stage Two (Optimistic Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	μ_0 (Case No.)	μ_1 (Case No.)	Wilcoxon p	$\mu_0 \neq \mu_1$	Clustered Rank
Native 500 ppi	59131.0	29.00	1	NA	NA	NA		1
Gaussian σ = 1.2730, r = 5, Odd-Even	37417.0	18.35	2	1	2	<0.0001	Yes	2
Gaussian σ = 1.3581, r = 5, Odd-Even	37338.0	18.31	3	2	3	0.9017		2
Gaussian σ = 1.1879, r = 5, Odd-Even	36515.0	17.91	4	2	4	0.0378		2
Gaussian σ = 1.4432, r = 5, Odd-Even	36423.0	17.86	5	2	5	0.0024		2
Gaussian σ = 1.1028, r = 4, Odd-Even	36227.5	17.77	6	2	6	0.0403		2
Gaussian σ = 1.5283, r = 6, Odd-Even	35440.0	17.38	7	2	7	<0.0001	Yes	3
Gaussian σ = 1.0177, r = 4, Odd-Even	35326.0	17.33	8	7	8	0.7248		3
Gaussian σ = 0.9326, r = 4, Odd-Even	34803.5	17.07	9	7	9	0.2302		3
Gaussian σ = 0.8475, r = 4, Odd-Even	34246.0	16.80	10	7	10	0.0645		3
Gaussian σ = 1.6134, r = 6, Odd-Even	32756.5	16.06	11	7	11	<0.0001	Yes	4
Gaussian σ = 0.7624, r = 3, Odd-Even	32663.5	16.02	12	11	12	0.7621		4
Gaussian σ = 0.6773, r = 3, Odd-Even	31352.0	15.38	13	11	13	0.0300		4
Gaussian σ = 1.6985, r = 6, Odd-Even	30921.5	15.17	14	11	14	<0.0001	Yes	5
Gaussian σ = 0.5922, r = 3, Odd-Even	29283.5	14.36	15	14	15	0.0058		5
Gaussian σ = 0.5642, r = 3, Odd-Even	28793.5	14.12	16	14	16	0.0010	Yes	6
Gaussian σ = 0.4220, r = 2, Even-Odd	27485.5	13.48	17	16	17	0.0006	Yes	7
Gaussian σ = 0.5071, r = 3, Odd-Odd	27038.0	13.26	18	17	18	0.6643		7
Gaussian σ = 0.3369, r = 2, Even-Odd	26947.0	13.22	19	17	19	0.0612		7
Averaging	26621.5	13.06	20	17	20	0.0687		7
Gaussian σ = 1.8687, r = 7, Odd-Even	26348.5	12.92	21	17	21	0.0411		7
Decimation, Even-Odd	25940.0	12.72	22	17	22	0.0002	Yes	8
Gaussian σ = 1.7836, r = 6, Odd-Odd	25220.0	12.37	23	22	23	0.4083		8
Transcoded B	24596.5	12.06	24	22	24	0.0031		8
Transcoded A	24571.0	12.05	25	22	25	0.0100		8
Spectral Truncation	23375.5	11.46	26	22	26	<0.0001	Yes	9
Gaussian σ = 1.9538, r = 7, Odd-Even	23269.0	11.41	27	26	27	0.8483		9
Gaussian σ = 2.0389, r = 7, Odd-Even	19999.0	9.81	28	26	28	<0.0001	Yes	10
Gaussian σ = 2.1240, r = 7, Odd-Even	16916.0	8.30	29	28	29	<0.0001	Yes	11

n	2039	
Friedman's statistic	11487.77	
DF	28	
p	<0.0001	(X2 approximation, corrected for ties)

In stage three of analysis, the Rank 4 (worst) decimation candidates from each sub-treatment were compared against each other as well as the cases that received no sub-treatment. The data from this stage of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 19). Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. As described in section 5.1.1, this comparison continued until a statistically significant difference was observed. The values that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered rank values was established, with each clustered rank representing data cases that are not significantly different from the others in the cluster. Each clustered group was then provided an ordinal clustered rank value.

Table 19 - Automated Matching Data Analysis, Stage Three (Conservative Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	μ_o (Case No.)	μ_t (Case No.)	Wilcoxon p	$\mu_o \neq \mu_t$	Clustered Rank
Native 500 ppi	59131.0	29.00	1	NA	NA	NA		1
Gaussian σ = 1.1879, r = 5, Even-Odd	37065.0	18.18	2	1	2	<0.0001	Yes	2
Gaussian σ = 1.1028, r = 4, Odd-Odd	36742.0	18.02	3	2	3	0.4044		2
Gaussian σ = 1.2730, r = 5, Even-Odd	36529.5	17.92	4	2	4	0.0958		2
Gaussian σ = 1.3581, r = 5, Even-Odd	36373.5	17.84	5	2	5	0.0146		2
Gaussian σ = 1.0177, r = 4, Odd-Odd	36061.0	17.69	6	2	6	0.1022		2
Gaussian σ = 0.9326, r = 4, Odd-Odd	35416.0	17.37	7	2	7	0.0016	Yes	3
Gaussian σ = 1.4432, r = 5, Even-Odd	35139.5	17.23	8	7	8	0.1463		3
Gaussian σ = 0.8475, r = 4, Odd-Odd	34016.0	16.68	9	7	9	0.0002	Yes	4
Gaussian σ = 1.5283, r = 6, Even-Odd	33613.0	16.49	10	9	10	0.1735		4
Gaussian σ = 0.7624, r = 3, Odd-Odd	33439.5	16.40	11	9	11	0.4342		4
Gaussian σ = 0.6773, r = 3, Even-Even	32740.5	16.06	12	9	12	0.0023		4
Gaussian σ = 1.6134, r = 6, Even-Odd	31678.0	15.54	13	9	13	<0.0001	Yes	5
Gaussian σ = 0.5071, r = 3, Odd-Even	30596.5	15.01	14	13	14	0.2406		5
Averaging	30417.5	14.92	15	13	15	0.1141		5
Gaussian σ = 0.5922, r = 3, Even-Odd	29691.0	14.56	16	13	16	0.0014	Yes	6
Gaussian σ = 1.6985, r = 6, Even-Odd	29469.0	14.45	17	16	17	0.3494		6
Gaussian σ = 0.5642, r = 3, Even-Odd	28953.0	14.20	18	16	18	0.0073		6
Transcoded A	28501.0	13.98	19	16	19	0.0073		6
Transcoded B	28208.5	13.83	20	16	20	0.0033		6
Spectral Truncation	26960.0	13.22	21	16	21	<0.0001		6
Gaussian σ = 1.7836, r = 6, Even-Odd	26817.5	13.15	22	21	22	0.6007		6
Gaussian σ = 0.4220, r = 2, Even-Even	25362.0	12.44	23	21	23	0.0012	Yes	7
Gaussian σ = 1.8687, r = 7, Even-Odd	24283.5	11.91	24	23	24	0.0298		7
Gaussian σ = 0.3369, r = 2, Even-Even	23056.5	11.31	25	23	25	<0.0001	Yes	8
NIST Decimation Even-Even	22149.0	10.86	26	25	26	0.0163		8
Gaussian σ = 1.9538, r = 7, Even-Odd	21112.0	10.35	27	25	27	0.0001	Yes	9
Gaussian σ = 2.0389, r = 7, Even-Odd	18217.5	8.93	28	27	28	<0.0001	Yes	10
Gaussian σ = 2.1240, r = 7, Even-Odd	15226.0	7.47	29	280	29	<0.0001	Yes	11

n	2039	
Friedman's statistic	12111.46	
DF	28	
p	<0.0001	(X2 approximation, corrected for ties)

5.3.2. Investigative Result 3

Based on rank analysis of the pared down cases, the Native 500 ppi Control Case yielded the best performance for the machine matcher among the methods utilizing the optimistic decimation strategy (stage 2 analysis). This was followed by several Gaussian methods. Post-hoc analysis using Wilcoxon indicates that the Native Control Case exhibits operational behavior that is statistically better than the Gaussian cases that follow it and can clearly be identified as the best. Post-hoc analysis also showed that the five Gaussian cases that followed the control case (Gaussian σ = 1.2730, r = 5, Odd-Even), (Gaussian σ = 1.3581, r = 5, Odd-Even), (Gaussian σ = 1.1879, r = 5, Odd-Even), (Gaussian σ = 1.4432, r = 5, Odd-Even) and (Gaussian σ = 1.1028, r = 4, Odd-Even) are equally as effective and have a score distribution that is not significantly different from each other and therefore these five methods share clustered rank after the Control Case as the next best performing case.

Analysis using the conservative decimation strategy (stage 3 analysis) showed that the Native 500 ppi Control Case again yields the best mean rank. Post-hoc analysis using Wilcoxon indicates that the Native Control Case exhibits operational behavior that is statistically better than the Gaussian cases that follow it and can clearly be identified as the best. Post-hoc analysis also showed that the five Gaussian cases that followed the Control Case (Gaussian σ = 1.1879, r = 5, Even-Odd), (Gaussian σ = 1.1028, r = 4, Odd-Odd), (Gaussian σ = 1.2730, r = 5, Even-Odd), (Gaussian σ

= 1.3581, r = 5, Even-Odd) and (Gaussian σ = 1.0177, r = 4, Odd-Odd) are equally as effective and have a score distribution that is not significantly different from each other and therefore these five methods share clustered rank after the Control Case as the next best performing case.

5.4. Investigative Goal 4: Signal Analysis and Assessment

5.4.1. Investigative Analysis 4

Frequency Spectrum Correlation Analysis

Data analysis by conducting a frequency spectrum correlation analysis using SIVV (see 3.2.3) yielded correlation scores potentially ranging from 0 to 1, with 1 representing 100% spectrum fidelity to the original sample. In the first phase of analysis, Friedman's statistical test was performed on each treatment that required the decimation sub-treatment. This resulted in Friedman's ranking of the sub-treatments (see Table 20). This allowed for stratification of the sub-treatments from best (highest mean rank) to worst (lowest mean rank), and provided an indication that at least one statistically significant difference in distributions exists between the four sub-treatments via Friedman's p. Mean ranks of the sub-treatments allows us to select the best sub-treatment for each treatment and propagate that to the next phase of analysis. In addition to selecting the top ranked sub-treatment, a Wilcoxon post-hoc analysis was performed between treatments at Rank 1 and at Rank 2 to determine if the top ranked sub-treatment is indeed different from that at Rank 2. Given a finding of no significant difference for this test, Wilcoxon comparison would have been made between treatment at Rank 1 with that at Rank 3 and Rank 4. For this metric, the Rank 1 candidate in all cases was shown to be significantly different than Rank 2, hence no further analysis for Rank 3 or Rank 4 was conducted.

Table 20 - Frequency Spectrum Correlation Analysis, Stage One (Sub-treatment Pare-Down, Adjusted α = 0.0167)

Treatment	Sub-treatment, Friedman's Mean Rank, Lower is Better								
	Even-Even	Even-Odd	Odd-Even	Odd-Odd	Friedman's p	X^2 statistic	p, Wilcoxon Rank 1 ≠ 2	p, Wilcoxon Rank 1 ≠ 3	p, Wilcoxon Rank 1 ≠ 4
Native 500 ppi	-	-	-	-	-	-	-	-	-
Averaging	-	-	-	-	-	-	-	-	-
NIST Spectral Truncation	-	-	-	-	-	-	-	-	-
Decimation	1.03	3.18	2.18	3.61	<0.0001	474.94	<0.0001	-	-
Transcoded A								-	-
Transcoded B								-	-
Gaussian σ = 0.3369, r = 2	1.03	3.17	2.20	3.61	<0.0001	469.70	<0.0001	-	-
Gaussian σ = 0.4220, r = 2	1.04	3.15	2.26	3.57	<0.0001	450.78	<0.0001	-	-
Gaussian σ = 0.5071, r = 3	1.17	2.97	2.36	3.51	<0.0001	362.50	<0.0001	-	-
Gaussian σ = 0.5642, r = 3	1.40	2.82	2.32	3.47	<0.0001	275.83	<0.0001	-	-
Gaussian σ = 0.5922, r = 3	1.50	2.74	2.32	3.44	<0.0001	236.83	<0.0001	-	-
Gaussian σ = 0.6773, r = 3	1.71	2.69	2.30	3.31	<0.0001	162.03	<0.0001	-	-
Gaussian σ = 0.7624, r = 3	1.93	2.54	2.32	3.22	<0.0001	106.18	<0.0001	-	-
Gaussian σ = 0.8475, r = 4	2.00	2.32	2.39	3.30	<0.0001	111.32	<0.0001	-	-
Gaussian σ = 0.9326, r = 4	1.96	2.28	2.41	3.36	<0.0001	129.76	<0.0001	-	-
Gaussian σ = 1.0177, r = 4	2.03	2.16	2.41	3.41	<0.0001	140.05	<0.0001	-	-
Gaussian σ = 1.1028, r = 4	2.02	2.02	2.51	3.46	<0.0001	167.06	<0.0001	-	-
Gaussian σ = 1.1879, r = 5	2.18	1.92	2.46	3.45	<0.0001	161.59	<0.0001	-	-
Gaussian σ = 1.2730, r = 5	2.23	1.85	2.55	3.38	<0.0001	151.57	<0.0001	-	-
Gaussian σ = 1.3581, r = 5	2.28	1.80	2.58	3.35	<0.0001	150.97	<0.0001	-	-
Gaussian σ = 1.4432, r = 5	2.34	1.77	2.55	3.35	<0.0001	152.95	<0.0001	-	-
Gaussian σ = 1.5283, r = 6	2.38	1.71	2.58	3.34	<0.0001	160.96	<0.0001	-	-
Gaussian σ = 1.6134, r = 6	2.38	1.70	2.53	3.40	<0.0001	176.95	<0.0001	-	-
Gaussian σ = 1.6985, r = 6	2.41	1.62	2.61	3.37	<0.0001	187.35	<0.0001	-	-
Gaussian σ = 1.7836, r = 6	2.46	1.58	2.62	3.35	<0.0001	191.15	<0.0001	-	-
Gaussian σ = 1.8687, r = 7	2.50	1.54	2.58	3.39	<0.0001	205.25	<0.0001	-	-
Gaussian σ = 1.9538, r = 7	2.53	1.48	2.64	3.36	<0.0001	215.00	<0.0001	-	-
Gaussian σ = 2.0389, r = 7	2.53	1.39	2.65	3.44	<0.0001	257.99	<0.0001	-	-
Gaussian σ = 2.1240, r = 7	2.56	1.35	2.67	3.43	<0.0001	265.07	<0.0001	-	-

In stage two of analysis, the treatments combined with the best (Rank 1) decimation sub-treatment were compared against each other as well as to the downsampling methods that included no decimation sub-treatment. The data from this phase of analysis was sorted according to mean rank, from largest (best) to smallest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (Table 21). Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. As described previously, this comparison continued until a statistically significant difference was observed. The cases that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered rank values was established, with each clustered rank representing data cases that are not significantly different from the others (and hence are clustered as such). Each clustered group was then provided an ordinal clustered rank value.

Table 21 - Frequency Spectrum Correlation Analysis, Stage Two (Optimistic Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	P_0 (Case No.)	P_1 (Case No.)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Native 500 ppi	5800.0	29.00	1	NA	NA	NA		1
Gaussian σ = 0.8475, r = 4, Odd-Odd	4351.0	21.76	2	1	2	<0.0001	Yes	2
Gaussian σ = 0.7624, r = 3, Odd-Odd	4336.0	21.68	3	2	3	0.1938		2
Gaussian σ = 0.5642, r = 3, Odd-Odd	4335.0	21.68	4	2	4	0.0326		2
Gaussian σ = 0.5922, r = 3, Odd-Odd	4330.0	21.65	5	2	5	0.0231		2
Gaussian σ = 0.6773, r = 3, Odd-Odd	4330.0	21.65	6	2	6	0.0524		2
Gaussian σ = 0.5071, r = 3, Odd-Odd	4242.0	21.21	7	2	7	0.0847		2
Gaussian σ = 0.9326, r = 4, Odd-Odd	4232.0	21.16	8	2	8	<0.0001	Yes	3
Gaussian σ = 0.4220, r = 2, Odd-Odd	4173.0	20.87	9	8	9	0.0779		3
Gaussian σ = 0.3369, r = 2, Odd-Odd	4070.0	20.35	10	8	10	0.2308		3
Gaussian σ = 1.0177, r = 4, Odd-Odd	4052.0	20.26	11	8	11	<0.0001	Yes	4
Decimation, Odd-Odd	3866.0	19.33	12	11	12	0.1347		4
Gaussian σ = 1.1028, r = 4, Odd-Odd	3852.0	19.26	13	11	13	<0.0001	Yes	5
Gaussian σ = 1.1879, r = 5, Odd-Odd	3581.0	17.91	14	13	14	<0.0001	Yes	6
Gaussian σ = 1.2730, r = 5, Odd-Odd	3270.0	16.35	15	14	15	<0.0001	Yes	7
Gaussian σ = 1.3581, r = 5, Odd-Odd	2980.0	14.90	16	15	16	<0.0001	Yes	8
Averaging	2761.0	13.81	17	16	17	0.5565		8
Gaussian σ = 1.4432, r = 5, Odd-Odd	2678.0	13.39	18	16	18	<0.0001	Yes	9
Spectral Truncation	2457.0	12.29	19	18	19	0.0382		9
Gaussian σ = 1.5283, r = 6, Odd-Odd	2353.0	11.77	20	18	20	<0.0001	Yes	10
Gaussian σ = 1.6134, r = 6, Odd-Odd	2015.0	10.08	21	20	21	<0.0001	Yes	11
Gaussian σ = 1.6985, r = 6, Odd-Odd	1684.0	8.42	22	21	22	<0.0001	Yes	12
Transcoded B	1501.0	7.51	23	22	23	0.0001	Yes	13
Transcoded A	1436.0	7.18	24	23	24	<0.0001	Yes	14
Gaussian σ = 1.7836, r = 6, Odd-Odd	1382.0	6.91	25	24	25	0.0255		14
Gaussian σ = 1.8687, r = 7, Odd-Odd	1112.0	5.56	26	24	26	0.7372		14
Gaussian σ = 1.9538, r = 7, Odd-Odd	857.0	4.29	27	24	27	0.0766		14
Gaussian σ = 2.0389, r = 7, Odd-Odd	603.0	3.02	28	24	28	<0.0001	Yes	15
Gaussian σ = 2.1240, r = 7, Odd-Odd	361.0	1.81	29	28	29	<0.0001	Yes	16

n	200	
Friedman's statistic	3956.95	
DF	28	
p	<0.0001	(X^2 approximation, corrected for ties)

In stage three of analysis, the Rank 4 (worst) decimation candidates from each sub-treatment were compared against each other as well as the cases that received no sub-treatment. The data from this stage of analysis was sorted according to mean rank, from largest (best) to smallest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 22) . Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. As described in section 5.1.1, this comparison continued until a statistically significant difference was observed. The cases that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a clustered list of rank values was established, with each clustered rank representing data cases that are not significantly different from the others in the cluster. Each clustered group was then provided an ordinal clustered rank value.

Table 22 - Frequency Spectrum Correlation Analysis, Stage Three (Conservative Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	P_0 (Case No.)	P_1 (Case No.)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Native 500 ppi	5800.0	29.00	1	NA	NA	NA		1
Gaussian σ = 0.7624 r = 3, Even-Even	4496.0	22.48	2	1	2	<0.0001	Yes	2
Gaussian σ = 0.8475 r = 4, Even-Even	4420.0	22.10	3	2	3	0.3998		2
Gaussian σ = 0.9326 r = 4, Even-Even	4320.0	21.60	4	3	4	0.0277		2
Gaussian σ = 0.6773 r = 3, Even-Even	4247.0	21.24	5	3	5	0.0001	Yes	3
Gaussian σ = 1.1879 r = 5, Even-Odd	4215.0	21.08	6	5	6	0.5426		3
Averaging	4207.0	21.04	7	5	7	0.4270		3
Gaussian σ = 1.0177 r = 4, Even-Even	4188.0	20.94	8	5	8	0.6605		3
Gaussian σ = 1.1028 r = 4, Even-Even	3993.0	19.97	9	5	9	0.4723		3
Gaussian σ = 1.2730 r = 5, Even-Odd	3917.0	19.59	10	5	10	0.0836		3
Gaussian σ = 0.5922 r = 3, Even-Even	3693.0	18.47	11	5	11	<0.0001	Yes	4
Gaussian σ = 1.3581 r = 5, Even-Odd	3608.0	18.04	12	11	12	0.8682		4
Spectral Truncation	3477.0	17.39	13	11	13	0.9942		4
Gaussian σ = 0.56419 r = 3, Even-Even	3371.0	16.86	14	11	14	<0.0001	Yes	5
Gaussian σ = 1.4432 r = 5, Even-Odd	3278.0	16.39	15	14	15	0.8615		5
Gaussian σ = 1.5283 r = 6, Even-Odd	2940.0	14.70	16	14	16	0.3284		5
Gaussian σ = 0.5071 r = 3, Even-Even	2817.0	14.09	17	14	17	<0.0001	Yes	6
Gaussian σ = 1.6134 r = 6, Even-Odd	2588.0	12.94	18	17	18	0.5738		6
Transcoded B	2383.0	11.92	19	17	19	<0.0001	Yes	7
Transcoded A	2323.0	11.62	20	19	20	<0.0001	Yes	8
Gaussian σ = 1.6985 r = 6, Even-Odd	2240.0	11.20	21	20	21	0.4242		8
Gaussian σ = 1.7836 r = 6, Even-Odd	1905.0	9.53	22	20	22	0.4046		8
Gaussian σ = 0.4220 r = 2, Even-Even	1780.0	8.90	23	20	23	<0.0001	Yes	9
Gaussian σ = 1.8687 r = 7, Even-Odd	1595.0	7.98	24	23	24	0.6015		9
Gaussian σ = 0.3369 r = 2, Even-Even	1301.0	6.51	25	23	25	<0.0001	Yes	10
Gaussian σ = 1.9538 r = 7, Even-Odd	1266.0	6.33	26	25	26	0.0345		10
Decimation-Only, Even-Even	1010.0	5.05	27	25	27	<0.0001	Yes	11
Gaussian σ = 2.0389 r = 7, Even-Odd	962.0	4.81	28	27	28	0.1725		11
Gaussian σ = 2.1240 r = 7, Even-Odd	660.0	3.30	29	27	29	0.4442		11
n	200							
Friedman's statistic	3353.31							
DF	28							
p	<0.0001	(X^2 approximation, corrected for ties)						

Root Mean Square Difference (RMSD) Analysis of Spectra

Analysis of the RMS difference between SIVV frequency spectra (see 3.2.3) yields values that may range from 0 to ∞, with 0 representing 100% fidelity to the frequency spectrum of the 500 ppi exemplar image. In the first phase of analysis, Friedman's statistical test was performed on each treatment that required the decimation sub-treatment. This resulted in Friedman's ranking of the sub-treatments (see Table 23). This allowed for stratification of the sub-treatments from best (lowest mean rank for RMSD) to worst (highest mean rank for RMSD), and provided an indication that at least one statistically significant difference in distributions exists between the four sub-treatments via Friedman's p. Friedman's ranking of the sub-treatments allows us to select the best sub-treatment for each treatment and propagate that to the next phase of analysis. In addition to selecting the top ranked sub-treatment per Friedman's analysis, a Wilcoxon post-hoc analysis was performed between Rank 1 and Rank 2 to determine if the top ranked sub-treatment is indeed different from Rank 2-4.

Table 23 - Root Mean Square Difference Analysis, Stage One (Sub-treatment Pare-Down, Adjusted α = 0.0167)

Treatment	Sub-treatment, Friedman's Mean Rank, Lower is Better								
	Even-Even	Even-Odd	Odd-Even	Odd-Odd	Friedman's p	X^2 statistic	p, Wilcoxon Rank 1 ≠ 2	p, Wilcoxon Rank 1 ≠ 3	p, Wilcoxon Rank 1 ≠ 4
Native 500 ppi	-	-	-	-	-	-	-	-	
Averaging	-	-	-	-	-	-	-	-	
NIST Spectral Truncation	-	-	-	-	-	-	-	-	
Decimation	3.94	1.86	2.75	1.47	<0.0001	432.78	<0.0001		
Transcoded A									
Transcoded B									
Gaussian σ = 0.3369, r = 2	3.94	1.90	2.72	1.45	<0.0001	429.68	<0.0001		
Gaussian σ = 0.4220, r = 2	3.91	1.91	2.68	1.52	<0.0001	399.47	<0.0001		
Gaussian σ = 0.5071, r = 3	3.74	2.00	2.71	1.56	<0.0001	326.72	<0.0001		
Gaussian σ = 0.5642, r = 3	3.56	2.15	2.75	1.55	<0.0001	263.77	<0.0001		
Gaussian σ = 0.5922, r = 3	3.46	2.27	2.75	1.52	<0.0001	239.69	<0.0001		
Gaussian σ = 0.6773, r = 3	3.20	2.61	2.65	1.56	<0.0001	168.97	<0.0001		
Gaussian σ = 0.7624, r = 3	2.80	2.89	2.56	1.76	<0.0001	94.37	<0.0001		
Gaussian σ = 0.8475, r = 4	2.54	3.04	2.48	1.95	<0.0001	70.84	<0.0001		
Gaussian σ = 0.9326, r = 4	2.36	3.07	2.50	2.08	<0.0001	62.33	<0.0001		
Gaussian σ = 1.0177, r = 4	2.24	3.13	2.54	2.10	<0.0001	74.69	<0.0001		
Gaussian σ = 1.1028, r = 4	2.16	3.30	2.39	2.16	<0.0001	105.45	0.0142		
Gaussian σ = 1.1879, r = 5	2.07	3.21	2.42	2.31	<0.0001	88.40	0.1591	0.0166	
Gaussian σ = 1.2730, r = 5	2.03	3.09	2.41	2.48	<0.0001	69.44	0.0072		
Gaussian σ = 1.3581, r = 5	2.09	3.06	2.37	2.49	<0.0001	60.35	0.0899	0.4349	<0.0001
Gaussian σ = 1.4432, r = 5	2.13	3.07	2.34	2.47	<0.0001	57.95	0.1344	0.3978	<0.0001
Gaussian σ = 1.5283, r = 6	2.15	3.06	2.42	2.37	<0.0001	55.13	0.3320	0.1049	<0.0001
Gaussian σ = 1.6134, r = 6	2.18	3.06	2.39	2.38	<0.0001	53.62	0.3236	0.3094	<0.0001
Gaussian σ = 1.6985, r = 6	2.21	3.06	2.35	2.39	<0.0001	52.41	0.4589	0.3903	<0.0001
Gaussian σ = 1.7836, r = 6	2.21	3.06	2.37	2.36	<0.0001	52.10	0.4671	0.2991	<0.0001
Gaussian σ = 1.8687, r = 7	2.23	3.06	2.45	2.27	<0.0001	53.63	0.3711	0.1395	<0.0001
Gaussian σ = 1.9538, r = 7	2.16	3.06	2.48	2.31	<0.0001	52.10	0.4284	0.0431	<0.0001
Gaussian σ = 2.0389, r = 7	2.20	3.06	2.44	2.31	<0.0001	53.79	0.5029	0.1988	<0.0001
Gaussian σ = 2.1240, r = 7	2.21	3.06	2.43	2.30	<0.0001	53.11	0.5655	0.0750	<0.0001

In stage two of analysis, the Rank 1 candidates from each sub-treatment were compared against each other as well as the cases that received no sub-treatment. The data from this phase of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 24). Post-hoc pairwise comparison was then performed down this list, comparing Case 1 to Case 2 and so on. As described above, this comparison continued until a statistically significant difference was observed. The cases that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered ranks was established, with each clustered rank representing cases that are not significantly different from the others (and hence are clustered as such). Each clustered group was then provided an ordinal clustered rank value.

Table 24 - Root Mean Square Difference Analysis, Stage Two (Optimistic Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	P_0 (Case No.)	P_1 (Case No.)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Native 500 ppi	200.0	1.00	1	NA	NA	NA		1
Gaussian σ = 0.5922, r = 3, Odd-Odd	1318.0	6.59	2	1	2	<0.0001	Yes	2
Gaussian σ = 0.6773, r = 3, Odd-Odd	1321.0	6.61	3	2	3	0.2347		2
Gaussian σ = 0.5642, r = 3, Odd-Odd	1369.0	6.85	4	2	4	0.0012	Yes	3
Gaussian σ = 0.7624, r = 3, Odd-Odd	1412.0	7.06	5	4	5	0.6756		3
Gaussian σ = 0.5071, r = 3, Odd-Odd	1473.0	7.37	6	4	6	<0.0001	Yes	4
Gaussian σ = 0.8475, r = 4, Odd-Odd	1576.0	7.88	7	6	7	0.2474		4
Gaussian σ = 0.4220, r = 2, Odd-Odd	1623.0	8.12	8	6	8	<0.0001	Yes	5
Gaussian σ = 0.3369, r = 2, Odd-Odd	1740.0	8.70	9	8	9	<0.0001	Yes	6
Gaussian σ = 0.9326, r = 4, Odd-Odd	1857.0	9.29	10	9	10	0.1642		6
Decimation, Odd-Odd	1923.0	9.62	11	9	11	<0.0001	Yes	7
Gaussian σ = 1.0177, r = 4, Odd-Odd	2194.0	10.97	12	11	12	0.0126		7
Spectral Truncation	2429.0	12.15	13	11	13	<0.0001	Yes	8
Gaussian σ = 1.1028, r = 4, Odd-Odd	2568.0	12.84	14	13	14	0.5830		8
Averaging	2673.0	13.37	15	13	15	0.0781		8
Gaussian σ = 1.1879, r = 5, Even-Even	2990.0	14.95	16	13	16	<0.0001	Yes	9
Gaussian σ = 1.2730, r = 5, Even-Even	3355.0	16.78	17	16	17	<0.0001	Yes	10
Transcoded B	3489.0	17.45	18	17	18	0.0556		10
Transcoded A	3589.0	17.95	19	17	19	0.0636		10
Gaussian σ = 1.3581, r = 5, Even-Even	3681.0	18.41	20	17	20	<0.0001	Yes	11
Gaussian σ = 1.4432, r = 5, Even-Even	3961.0	19.81	21	20	21	<0.0001	Yes	12
Gaussian σ = 1.5283, r = 6, Even-Even	4215.0	21.08	22	21	22	<0.0001	Yes	13
Gaussian σ = 1.6134, r = 6, Even-Even	4460.0	22.30	23	22	23	<0.0001	Yes	14
Gaussian σ = 1.6985, r = 6, Even-Even	4703.0	23.52	24	23	24	<0.0001	Yes	15
Gaussian σ = 1.7836, r = 6, Even-Even	4938.0	24.69	25	24	25	<0.0001	Yes	16
Gaussian σ = 1.8687, r = 7, Even-Even	5160.0	25.80	26	25	26	<0.0001	Yes	17
Gaussian σ = 1.9538, r = 7, Even-Even	5384.0	26.92	27	26	27	<0.0001	Yes	18
Gaussian σ = 2.0389, r = 7, Even-Even	5599.0	28.00	28	27	28	<0.0001	Yes	19
Gaussian σ = 2.1240, r = 7, Even-Even	5800.0	29.00	29	28	29	<0.0001	Yes	20
n	200							
Friedman's statistic	4659.04							
DF	28							
p	<0.0001	(X2 approximation, corrected for ties)						

In stage three of analysis, the Rank 4 (worst) decimation candidates from each sub-treatment were compared against each other as well as the cases that received no sub-treatment. The data from this stage of analysis was sorted according to mean rank, from smallest (best) to largest (worst) and an ordinal case number ranging from 1 to 29 was assigned to this sorted list (see Table 25). Post-hoc analysis was then performed down this list, comparing Case 1 to Case 2 and so on. As above, this comparison continued until a statistically significant difference was observed. The cases that did not demonstrate a statistically significant difference were clustered into one singular rank. This process continued until a list of clustered ranks was established, with each clustered rank representing data that is not significantly different from the others (and hence are clustered as such). Each clustered group was then provided an ordinal clustered rank value.

Table 25 - Root Mean Square Difference Analysis, Stage Three (Conservative Decimation Pare-Down, Adjusted α = 0.0018)

Treatment	Rank sum	Mean Rank	Case No.	P_0 (Case No)	P_1 (Case No)	Wilcoxon p	$P_0 \neq P_1$	Clustered Rank
Native 500 ppi	200.0	1.00	1	NA	NA	NA		1
Gaussian σ = 0.7624, r = 3, Even-Odd	1086.0	5.43	2	1	2	<0.0001	Yes	2
Gaussian σ = 0.6773, r = 3, Even-Even	1242.0	6.21	3	2	3	0.0011	Yes	3
Averaging	1287.0	6.44	4	3	4	<0.0001	Yes	4
Gaussian σ = 0.8475, r = 4, Even-Odd	1317.0	6.59	5	4	5	0.1192		4
Spectral Truncation	1366.0	6.83	6	4	6	0.0781		4
Gaussian σ = 0.9326, r = 4, Even-Odd	1622.0	8.11	7	4	7	0.5402		4
Gaussian σ = 0.5922, r = 3, Even-Even	1679.0	8.40	8	4	8	<0.0001	Yes	5
Gaussian σ = 0.5642, r = 3, Even-Even	1949.0	9.75	9	8	9	<0.0001	Yes	6
Gaussian σ = 1.0177, r = 4, Even-Odd	1970.0	9.85	10	9	10	0.6998		6
Gaussian σ = 1.1028, r = 4, Even-Odd	2359.0	11.80	11	9	11	0.0285		6
Transcoded B	2487.0	12.44	12	9	12	<0.0001	Yes	7
Gaussian σ = 0.5071, r = 3, Even-Even	2572.0	12.86	13	12	13	0.0468		7
Transcoded A	2595.0	12.98	14	12	14	<0.0001	Yes	8
Gaussian σ = 1.1879, r = 5, Even-Odd	2738.0	13.69	15	14	15	0.0128		8
Gaussian σ = 1.2730, r = 5, Even-Odd	3095.0	15.48	16	14	16	<0.0001	Yes	9
Gaussian σ = 0.4220, r = 2, Even-Even	3230.0	16.15	17	16	17	0.1748		9
Gaussian σ = 1.3581, r = 5, Even-Odd	3415.0	17.08	18	16	18	<0.0001	Yes	10
Gaussian σ = 0.3369, r = 2, Even-Even	3601.0	18.01	19	18	19	0.1113		10
Gaussian σ = 1.4432, r = 5, Even-Odd	3725.0	18.63	20	18	20	<0.0001	Yes	11
Decimation, Even-Even	3865.0	19.33	21	20	21	0.0011	Yes	12
Gaussian σ = 1.5283, r = 6, Even-Odd	4007.0	20.04	22	21	22	<0.0001	Yes	13
Gaussian σ = 1.6134, r = 6, Even-Odd	4291.0	21.46	23	22	23	<0.0001	Yes	14
Gaussian σ = 1.6985, r = 6, Even-Odd	4588.0	22.94	24	23	24	<0.0001	Yes	15
Gaussian σ = 1.7836, r = 6, Even-Odd	4857.0	24.29	25	24	25	<0.0001	Yes	16
Gaussian σ = 1.8687, r = 7, Even-Odd	5121.0	25.61	26	25	26	<0.0001	Yes	17
Gaussian σ = 1.9538, r = 7, Even-Odd	5357.0	26.79	27	26	27	<0.0001	Yes	18
Gaussian σ = 2.0389, r = 7, Even-Odd	5584.0	27.92	28	27	28	<0.0001	Yes	19
Gaussian σ = 2.1240, r = 7, Even-Odd	5795.0	28.98	29	28	29	<0.0001	Yes	20
n	200							
Friedman's statistic	4482.30							
DF	28							
p	<0.0001	(X2 approximation, corrected for ties)						

5.4.2. Investigative Result 4

Using data obtained by frequency spectrum correlation analysis, Friedman's test identifies the native 500 ppi control (Case No.1) as the best at Rank 1 among the cases based on an optimistic decimation strategy (stage 2 analysis). Following the Control Case were six Gaussian cases of (Gaussian σ = 0.8475, r = 4, Odd-Odd), (Gaussian σ = 0.7624, r = 3, Odd-Odd), (Gaussian σ = 0.5642, r = 3, Odd-Odd), (Gaussian σ = 0.5922, r = 3, Odd-Odd), (Gaussian σ = 0.6773, r = 3, Odd-Odd) and (Gaussian σ = 0.5071, r = 3, Odd-Odd). Post-hoc analysis indicated the Rank 1 control (Case No.1) is different from Case No.2, but the difference between Case No.2 through Case No.7 was not statistically significant and are clustered as Rank 2. For cases based on a conservative decimation strategy (stage 3 analysis), Friedman's test identifies the native 500 ppi control (Case No.1) as the best at Rank 1. Following the control cases were 3 Gaussian cases of (Gaussian σ = 0.7624 r = 3, Even-Even), (Gaussian σ = 0.8475 r = 4, Even-Even) and (Gaussian σ = 0.9326 r = 4, Even-Even). Post-hoc analysis indicated the Rank 1 control (Case No.1) is statistically different from Case No.2, but the difference between Case No.2 through Case No.4 was not statistically significant and they are therefore clustered as Rank 2.

Using data obtained by RMSD analysis, Friedman's test again identifies the Native 500 ppi Control (Case No.1) as the best at Rank 1 among the cases based on an optimistic decimation strategy (stage 2 analysis). Following the control cases were the two Gaussian cases of (Gaussian σ = 0.5922, r = 3, Odd-Odd) and (Gaussian σ = 0.6773, r = 3, Odd-Odd). Post-hoc analysis indicated the Rank 1 control (Case No.1) is different from Case No.2, but the difference between Case No.2 and Case No.3 was not statistically significant and are clustered as Rank 2. For cases based on a conservative decimation strategy (stage 3 analysis), Friedman's test identifies the native 500 ppi control (Case No.1) as the best at rank-1. Following the control is the Gaussian case of (Gaussian σ = 0.7624, r = 3, Even-Odd). Post-hoc analysis indicated the Rank 1 control (Case No.1) is statistically different from Case No.2, and Case No.2 is statistically different from Case No.3.

5.5. Investigate Goal 5: Optimal Strategy Determination

Several approaches were explored with the goal of optimizing a process given multiple success measures across multiple treatments. Many of these approaches were geared towards a winner-take-all strategy or were not suitable in cases where different treatments and success measures yielded similar behaviors that may not consistently stratify operational behavior and result in ambiguous results. In this section we explore two such methods and present those findings.

5.5.1. Clustered Rank Summation Analysis

In the process of analyzing the data, the results of all experimental treatments were first stratified using a clustered rank summation strategy (See 5.1 - 5.4). Next, the clustered ranks for each metric across all treatments were summed, and the optimal strategy is represented by the minima of the sum of all clustered rank for each treatment.

Table 26 shows the clustered rank summation analysis of the optimistic decimation cases yielding an optimistic gauge of performance, while Table 27 shows the clustered rank summation analysis of the conservative decimation cases yielding a conservative gauge of performance. In both approaches, the Native 500 ppi Control Case emerges as the best all-around using all performance measures and metrics.

The treatment cases that expert examiners gauged as best are marked by a dashed border.

Table 26 - Clustered Rank Summation Using Optimistic Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Clustered Ranks (lower is better)	
Native 500 ppi	2	1	1	1	5	BEST (Control)
Gaussian σ = 0.6773, r = 3	1	4	2	2	9	
Gaussian σ = 0.7624, r = 3	1	4	2	3	10	
Gaussian σ = 0.5922, r = 3	1	5	2	2	10	
Gaussian σ = 0.8475, r = 4	1	3	2	4	10	
Gaussian σ = 0.5642, r = 3	1	6	2	3	12	
Gaussian σ = 0.9326, r = 4	1	3	3	6	13	
Gaussian σ = 0.5071, r = 3	1	7	2	4	14	
Gaussian σ = 1.0177, r = 4	1	3	4	7	15	
Gaussian σ = 0.4220, r = 2	1	7	3	5	16	
Gaussian σ = 1.1028, r = 4	2	2	5	8	17	
Gaussian σ = 0.3369, r = 2	1	7	3	6	17	
Gaussian σ = 1.1879, r = 5	3	2	6	9	20	
Gaussian σ = 1.2730, r = 5	3	2	7	10	22	
Decimation	3	8	4	7	22	
Gaussian σ = 1.3581, r = 5	4	2	8	11	25	
Gaussian σ = 1.4432, r = 5	4	2	9	12	27	
Averaging	4	7	8	8	27	
Spectral Truncation	2	9	9	8	28	
Gaussian σ = 1.5283, r = 6	4	3	10	13	30	
Gaussian σ = 1.6134, r = 6	4	4	11	14	33	
Transcoded B	3	8	13	10	34	
Transcoded A	4	8	14	10	36	
Gaussian σ = 1.6985, r = 6	4	5	12	15	36	
Gaussian σ = 1.7836, r = 6	5	8	14	16	43	
Gaussian σ = 1.8687, r = 7	5	7	14	17	43	
Gaussian σ = 1.9538, r = 7	5	9	14	18	46	
Gaussian σ = 2.0389, r = 7	5	10	15	19	49	
Gaussian σ = 2.1240, r = 7	6	11	16	20	53	

Table 27 - Clustered Rank Summation Using Conservative Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Clustered Ranks (lower is better)	
Native 500 ppi	1	1	1	1	4	BEST (Control)
Gaussian σ = 0.9326, r = 4	1	3	2	4	10	
Gaussian σ = 0.7624, r = 3	2	4	2	2	10	
Gaussian σ = 0.8475, r = 4	1	4	2	4	11	
Gaussian σ = 0.6773, r = 3	1	4	3	3	11	
Gaussian σ = 1.0177, r = 4	1	2	3	6	12	
Gaussian σ = 1.1028, r = 4	2	2	3	6	13	
Gaussian σ = 1.1879, r = 5	2	2	3	8	15	
Spectral Truncation	1	6	4	4	15	
Averaging	3	5	3	4	15	
Gaussian σ = 0.5922, r = 3	2	6	4	5	17	
Gaussian σ = 1.2730, r = 5	4	2	3	9	18	
Gaussian σ = 0.5642, r = 3	2	6	5	6	19	
Gaussian σ = 1.3581, r = 5	4	2	4	10	20	
Gaussian σ = 0.5071, r = 3	3	5	6	7	21	
Transcoded B	2	6	7	7	22	
Gaussian σ = 1.4432, r = 5	5	3	5	11	24	
Transcoded A	4	6	8	8	26	
Gaussian σ = 1.5283, r = 6	5	4	5	13	27	
Gaussian σ = 0.4220, r = 2	3	7	9	9	28	
Gaussian σ = 1.6134, r = 6	6	5	6	14	31	
Gaussian σ = 0.3369, r = 2	4	8	10	10	32	
Gaussian σ = 1.6985, r = 6	6	6	8	15	35	
Decimation	6	8	11	12	37	
Gaussian σ = 1.7836, r = 6	7	6	8	16	37	
Gaussian σ = 1.8687, r = 7	7	7	9	17	40	
Gaussian σ = 1.9538, r = 7	8	9	10	18	45	
Gaussian σ = 2.0389, r = 7	8	10	11	19	48	
Gaussian σ = 2.1240, r = 7	8	11	11	20	50	

5.5.2. Ordinal Rank Summation Analysis

A more conservative approach in selecting an optimal strategy is to add the ordinal rankings of each treatment together. Next the ordinal rankings for all treatments were summed, and the optimal strategy is represented by the minima of the sum of all ranks for each treatment. Since this method utilizes integer ranks, summation of results for each treatment may yield a higher probability of ties.

Table 28 shows the mean rank summation analysis of the optimistic decimation cases yielding an optimistic gauge of performance, while Table 29 shows the mean rank summation analysis of the conservative decimation cases yielding a conservative gauge of performance.

The case ranked best by expert examiners is emphasized by a dashed rectangle in the table. This case is traditionally considered operational ground truth.

Table 28 - Ordinal Rank Summation Using Optimistic Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Mean Ranks (lower is better)	
Native 500 ppi	11	1	1	1	14	BEST (Control)
Gaussian σ = 0.8475, r = 4	3	10	2	7	22	
Gaussian σ = 0.5922, r = 3	1	15	5	2	23	
Gaussian σ = 0.7624, r = 3	5	12	3	5	25	
Gaussian σ = 0.6773, r = 3	4	13	6	3	26	
Gaussian σ = 0.5642, r = 3	8	16	4	4	32	
Gaussian σ = 0.9326, r = 4	6	9	8	10	33	
Gaussian σ = 0.5071, r = 3	2	18	7	6	33	
Gaussian σ = 1.0177, r = 4	7	8	11	12	38	
Gaussian σ = 0.4220, r = 2	10	17	9	8	44	
Gaussian σ = 1.1028, r = 4	12	6	13	14	45	
Gaussian σ = 0.3369, r = 2	9	19	10	9	47	
Gaussian σ = 1.1879, r = 5	14	4	14	16	48	
Gaussian σ = 1.2730, r = 5	16	2	15	17	50	
Gaussian σ = 1.3581, r = 5	18	3	16	20	57	
Decimation	15	22	12	11	60	
Gaussian σ = 1.4432, r = 5	19	5	18	21	63	
Gaussian σ = 1.5283, r = 6	21	7	20	22	70	
Spectral Truncation	13	26	19	13	71	
Averaging	20	20	17	15	72	
Gaussian σ = 1.6134, r = 6	22	11	21	23	77	
Transcoded B	17	24	23	18	82	
Gaussian σ = 1.6985, r = 6	24	14	22	24	84	
Transcoded A	23	25	24	19	91	
Gaussian σ = 1.7836, r = 6	25	23	25	25	98	
Gaussian σ = 1.8687, r = 7	26	21	26	26	99	
Gaussian σ = 1.9538, r = 7	27	27	27	27	108	
Gaussian σ = 2.0389, r = 7	28	28	28	28	112	
Gaussian σ = 2.1240, r = 7	29	29	29	29	116	

Table 29 - Ordinal Rank Summation Using Conservative Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Mean Ranks (lower is better)	
Native 500 ppi	2	1	1	1	5	BEST (Control)
Gaussian σ = 0.9326, r = 4	3	7	4	7	21	
Gaussian σ = 0.8475, r = 4	5	9	3	5	22	
Gaussian σ = 0.7624, r = 3	7	11	2	2	22	
Gaussian σ = 0.6773, r = 3	6	12	5	3	26	
Gaussian σ = 1.0177, r = 4	4	6	8	10	28	
Gaussian σ = 1.1028, r = 4	9	3	9	11	32	
Gaussian σ = 1.1879, r = 5	12	2	6	15	35	
Spectral Truncation	1	21	13	6	41	
Averaging	15	15	7	4	41	
Gaussian σ = 0.5922, r = 3	10	16	11	8	45	
Gaussian σ = 1.2730, r = 5	16	4	10	16	46	
Gaussian σ = 0.5642, r = 3	11	18	14	9	52	
Gaussian σ = 1.3581, r = 5	18	5	12	18	53	
Gaussian σ = 0.5071, r = 3	13	14	17	13	57	
Transcoded B	8	20	19	12	59	
Gaussian σ = 1.4432, r = 5	20	8	15	20	63	
Gaussian σ = 1.5283, r = 6	21	10	16	22	69	
Transcoded A	19	19	20	14	72	
Gaussian σ = 1.6134, r = 6	22	13	18	23	76	
Gaussian σ = 0.4220, r = 2	14	23	23	17	77	
Gaussian σ = 1.6985, r = 6	24	17	21	24	86	
Gaussian σ = 0.3369, r = 2	17	25	25	19	86	
Gaussian σ = 1.7836, r = 6	25	22	22	25	94	
Decimation	23	26	27	21	97	
Gaussian σ = 1.8687, r = 7	26	24	24	26	100	
Gaussian σ = 1.9538, r = 7	27	27	26	27	107	
Gaussian σ = 2.0389, r = 7	28	28	28	28	112	
Gaussian σ = 2.1240, r = 7	29	29	29	29	116	

5.5.3. Mean Rank Summation Analysis

An even more conservative and sensitive approach in selecting an optimal strategy is to simply add the mean rankings of each treatment together. Next the mean ranking for all treatment measures were summed, and the optimal strategy is represented by the minima of the sum of all ranks for each treatment. Since this method utilizes mean ranks, summation of results for each treatment has a lesser chance of yielding ties and can more easily stratify treatments.

Table 30 shows the mean rank summation analysis of the optimistic decimation cases yielding an optimistic gauge of performance, while Table 31 shows the mean rank summation analysis of the conservative decimation cases yielding a conservative gauge of performance.

The case ranked best by expert examiners is emphasized by a dashed rectangle in the table. This case is traditionally considered operational ground truth.

Table 30 - Mean Rank Summation Using Optimistic Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Mean Ranks (lower is better)	
Native 500 ppi	11.07	1.00	1.00	1.00	14.07	BEST (Control)
Gaussian σ = 0.8475, r = 4	9.16	13.20	8.25	7.88	38.49	
Gaussian σ = 0.6773, r = 3	9.27	14.62	8.35	6.61	38.85	
Gaussian σ = 0.7624, r = 3	9.81	13.98	8.32	7.06	39.17	
Gaussian σ = 0.5922, r = 3	8.63	15.64	8.35	6.59	39.21	
Gaussian σ = 0.9326, r = 4	9.85	12.93	8.84	9.29	40.90	
Gaussian σ = 0.5642, r = 3	9.93	15.88	8.33	6.85	40.98	
Gaussian σ = 0.5071, r = 3	9.10	16.74	8.79	7.37	41.99	
Gaussian σ = 1.0177, r = 4	9.88	12.67	9.74	10.97	43.26	
Gaussian σ = 0.4220, r = 2	10.32	16.52	9.14	8.12	44.09	
Gaussian σ = 0.3369, r = 2	10.05	16.78	9.65	8.70	45.18	
Gaussian σ = 1.1028, r = 4	11.47	12.23	10.74	12.84	47.29	
Decimation	14.61	17.28	10.67	9.62	52.18	
Gaussian σ = 1.1879, r = 5	13.66	12.09	12.10	14.95	52.79	
Gaussian σ = 1.2730, r = 5	15.16	11.65	13.65	16.78	57.24	
Spectral Truncation	11.53	18.54	17.72	12.15	59.92	
Gaussian σ = 1.3581, r = 5	16.95	11.69	15.10	18.41	62.14	
Averaging	18.14	16.94	16.20	13.37	64.64	
Gaussian σ = 1.4432, r = 5	17.90	12.14	16.61	19.81	66.45	
Gaussian σ = 1.5283, r = 6	18.28	12.62	18.24	21.08	70.20	
Transcoded B	15.19	17.94	22.50	17.45	73.07	
Gaussian σ = 1.6134, r = 6	19.79	13.94	19.93	22.30	75.95	
Transcoded A	20.23	17.95	22.82	17.95	78.94	
Gaussian σ = 1.6985, r = 6	20.61	14.83	21.58	23.52	80.54	
Gaussian σ = 1.7836, r = 6	22.02	17.63	23.09	24.69	87.43	
Gaussian σ = 1.8687, r = 7	22.13	17.08	24.44	25.80	89.45	
Gaussian σ = 1.9538, r = 7	22.79	18.59	25.72	26.92	94.02	
Gaussian σ = 2.0389, r = 7	23.56	20.19	26.99	28.00	98.73	
Gaussian σ = 2.1240, r = 7	23.94	21.70	28.20	29.00	102.84	

Table 31 - Mean Rank Summation Using Conservative Decimation Strategy

Treatment	Expert Examiner	Automated Matching	Frequency Spectrum Correlation	Root Mean Square Difference	Sum of Mean Ranks (lower is better)	
Native 500 ppi	7.66	1.00	1.00	1.00	10.66	BEST (Control)
Gaussian σ = 0.7624, r = 3	10.26	13.60	7.52	5.43	36.81	
Gaussian σ = 0.8475, r = 4	9.66	13.32	7.90	6.59	37.46	
Gaussian σ = 0.9326, r = 4	9.24	12.63	8.40	8.11	38.38	
Gaussian σ = 0.6773, r = 3	9.97	13.94	8.77	6.21	38.89	
Gaussian σ = 1.0177, r = 4	9.29	12.31	9.06	9.85	40.51	
Spectral Truncation	7.57	16.78	12.62	6.83	43.79	
Gaussian σ = 1.1028, r = 4	10.43	11.98	10.04	11.80	44.24	
Averaging	14.16	15.08	8.97	6.44	44.64	
Gaussian σ = 1.1879, r = 5	11.73	11.82	8.93	13.69	46.17	
Gaussian σ = 0.5922, r = 3	10.91	15.44	11.54	8.40	46.28	
Gaussian σ = 0.5642, r = 3	11.52	15.80	13.15	9.75	50.21	
Gaussian σ = 1.2730, r = 5	15.08	12.08	10.42	15.48	53.05	
Gaussian σ = 0.5071, r = 3	13.07	14.99	15.92	12.86	56.83	
Transcoded B	10.30	16.17	18.09	12.44	56.99	
Gaussian σ = 1.3581, r = 5	16.22	12.16	11.96	17.08	57.41	
Gaussian σ = 1.4432, r = 5	17.62	12.77	13.61	18.63	62.62	
Transcoded A	16.87	16.02	18.39	12.98	64.25	
Gaussian σ = 1.5283, r = 6	18.78	13.51	15.30	20.04	67.62	
Gaussian σ = 0.4220, r = 2	13.93	17.56	21.10	16.15	68.74	
Gaussian σ = 1.6134, r = 6	19.67	14.46	17.06	21.46	72.64	
Gaussian σ = 0.3369, r = 2	15.70	18.69	23.50	18.01	75.89	
Gaussian σ = 1.6985, r = 6	21.08	15.55	18.80	22.94	78.37	
Decimation	19.81	19.14	24.95	19.33	83.22	
Gaussian σ = 1.7836, r = 6	21.65	16.85	20.48	24.29	83.26	
Gaussian σ = 1.8687, r = 7	22.80	18.09	22.03	25.61	88.52	
Gaussian σ = 1.9538, r = 7	22.97	19.65	23.67	26.79	93.07	
Gaussian σ = 2.0389, r = 7	23.31	21.07	25.19	27.92	97.48	
Gaussian σ = 2.1240, r = 7	23.80	22.53	26.70	28.98	102.01	

5.5.4. Mean Rank Summation Analysis for Expert Examiner Winner-Take-All

Since examiner opinion is widely recognized as the ground truth in fingerprint examination processing pathways, Table 32 below shows expert examiner mean rank analysis as a product of the mean rank for both the optimistic and conservative decimation strategies. The Native 500 ppi Control Case emerged as the best, while Gaussian (Gaussian σ = 0.8475, r = 4) emerged as the next best, followed by (Gaussian σ = 0.9326, r = 4) and Spectral Truncation.

Table 32 - Expert Examiner Winner-Take-All Mean Rank Analysis

Treatment	Optimistic Decimation Mean Rank	Conservative Decimation Mean Rank	Mean Rank Summation
Native 500 ppi	11.07	7.66	18.72
Gaussian σ = 0.8475, r = 4	9.16	9.66	18.82
Gaussian σ = 0.9326, r = 4	9.85	9.24	19.08
Spectral Truncation	11.53	7.57	19.09
Gaussian σ = 1.0177, r = 4	9.88	9.29	19.17
Gaussian σ = 0.6773, r = 3	9.27	9.97	19.24
Gaussian σ = 0.5922, r = 3	8.63	10.91	19.54
Gaussian σ = 0.7624, r = 3	9.81	10.26	20.07
Gaussian σ = 0.5642, r = 3	9.93	11.52	21.45
Gaussian σ = 1.1028, r = 4	11.47	10.43	21.90
Gaussian σ = 0.5071, r = 3	9.10	13.07	22.17
Gaussian σ = 0.4220, r = 2	10.32	13.93	24.24
Gaussian σ = 1.1879, r = 5	13.66	11.73	25.39
Transcoded B	15.19	10.30	25.49
Gaussian σ = 0.3369, r = 2	10.05	15.70	25.75
Gaussian σ = 1.2730, r = 5	15.16	15.08	30.24
Averaging	18.14	14.16	32.30
Gaussian σ = 1.3581, r = 5	16.95	16.22	33.16
Decimation	14.61	19.81	34.43
Gaussian σ = 1.4432, r = 5	17.90	17.62	35.52
Gaussian σ = 1.5283, r = 6	18.28	18.78	37.05
Transcoded A	20.23	16.87	37.10
Gaussian σ = 1.6134, r = 6	19.79	19.67	39.45
Gaussian σ = 1.6985, r = 6	20.61	21.08	41.69
Gaussian σ = 1.7836, r = 6	22.02	21.65	43.67
Gaussian σ = 1.8687, r = 7	22.13	22.80	44.93
Gaussian σ = 1.9538, r = 7	22.79	22.97	45.76
Gaussian σ = 2.0389, r = 7	23.56	23.31	46.87
Gaussian σ = 2.1240, r = 7	23.94	23.80	47.74

5.6. Uncertainty Measurement

We compare the various downsampling methods via the ordered mean rank scores output from the Friedman's tests applied to the measurements as there is no conventional statistic by which to assess the uncertainty or standard error of the mean rank scores. However, we are able to provide an estimate of uncertainty using a bootstrap procedure [WU1], [WU2], [WU3]. We first generate a set of mean rank scores by applying the Friedman test to each set of measurements. For the methods requiring decimation, the best decimation case was selected (sub-treatment pare-down, as described in 4.2) prior to conducting the Friedman test. We sort these mean rank scores in either ascending or descending order depending upon the nature of the particular metric, i.e., whether lowest or highest score is "best". We sort columns of the data matrix in the same order and resample randomly with replacement to generate 500 replicates of the data. We apply the Friedman's test to each of the data replicates to generate 2000 sets (vectors) of the measure of interest (such as sample mean or median) and assemble these into a matrix. Each column of the matrix represents a distribution of that measure for each of the *k* downsampling methods. The 95 % confidence intervals for the measures are then taken as the 0.025^{th} and 0.975^{th} quantiles of the distribution of mean ranks. This bootstrap procedure was applied to sample mean and sample median and to the Friedman's mean rank for each of the metrics examined in this study (see Table 4).

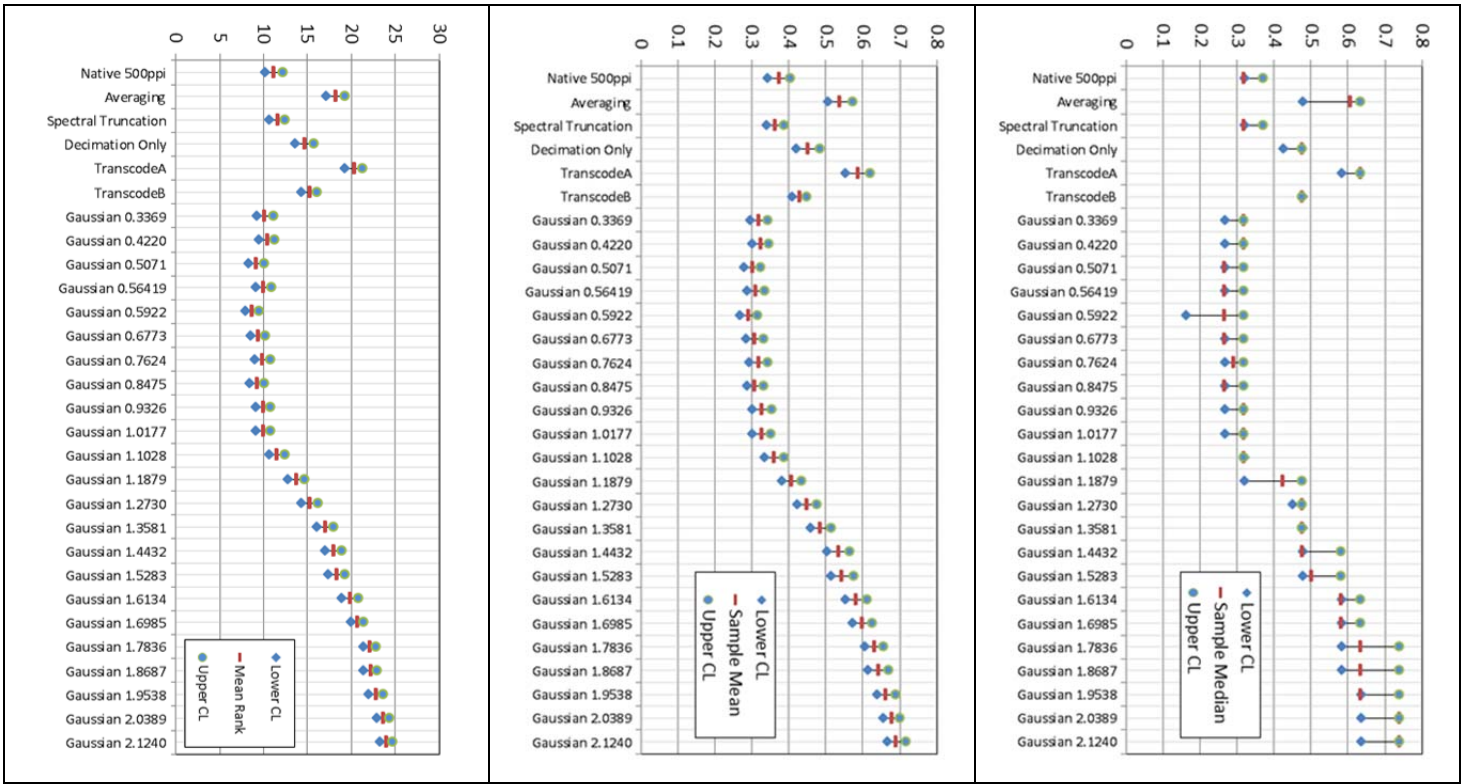
5.6.1. Expert Examiner Subjective Assessment Uncertainty

Figure 14 shows 95 % Confidence Intervals for Expert Examiner Subjective assessments based on 2000 replicates of Friedman's Mean Rank analysis. While this bootstrap analysis can be performed on the Friedman's mean rank, it cannot easily be performed on the sample mean and sample median. The reason for this is that while the scores for the expert examiner analysis portion of this study take on numeric form (111, 222, etc.), these scores are actually classifiers of the observations from the examiners, do not form a continuous scale and therefore examination of the means or medians of these classifiers has no meaning and cannot be used to formulate a conclusions. The classifiers on

the other hand can be ordered and lend themselves well to ranking tests and analysis such as Friedman's. To mitigate this shortfall, the classifier decisions were assigned a relative score. The relative score provides a fractional scale that ranges from 0 (best) to 1 (worst) and uniformly distributed to the all possible classification decisions as shown below in Table 33. Bootstrap analysis was then performed on sample mean and sample median of these values and the results are shown in Figure 14.

Table 33 - Classifier Score Assignments

Case Number	Decision Classifier	Assigned Score
1	111 (best)	0.0000
2	112	0.0526
3	113	0.1053
4	122	0.1579
5	114	0.2105
6	123	0.2632
7	222	0.3158
8	124	0.3684
9	133	0.4211
10	223	0.4737
11	134	0.5263
12	224	0.5789
13	233	0.6316
14	144	0.6842
15	234	0.7368
16	333	0.7895
17	244	0.8421
18	334	0.8947
19	344	0.9474
20	444 (worst)	1.0000

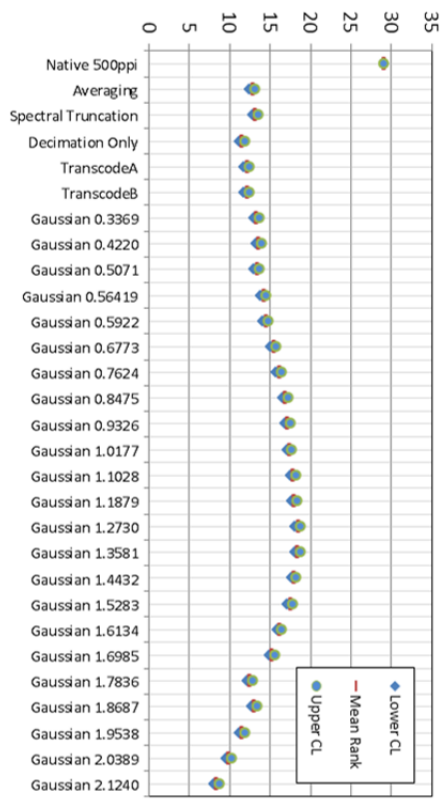


Mean Rank (Range 1-29, Lower is Better) | Sample Mean (Range 0 to 1, Lower is Better) | Sample Median (Range 0 to 1, Lower is Better)

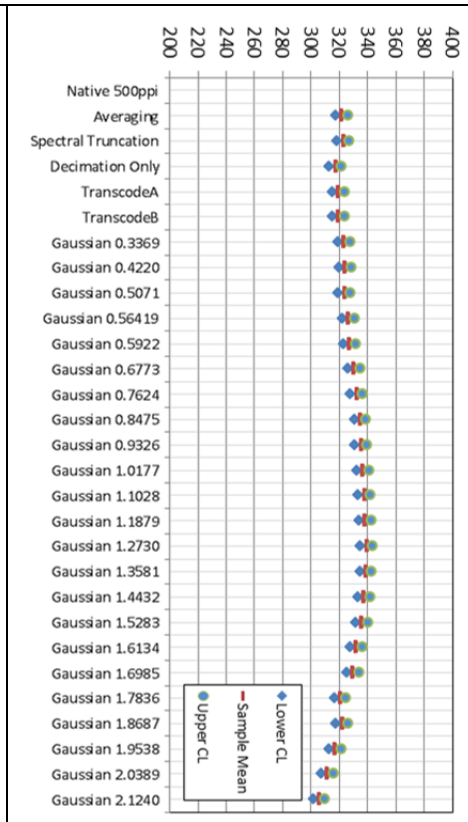
Figure 14 - Expert Examiner Subjective Assessment 95 % Confidence Intervals for Mean Rank, Mean and Median using Bootstrap Analysis

5.6.2. Automated Matching Assessment Uncertainty

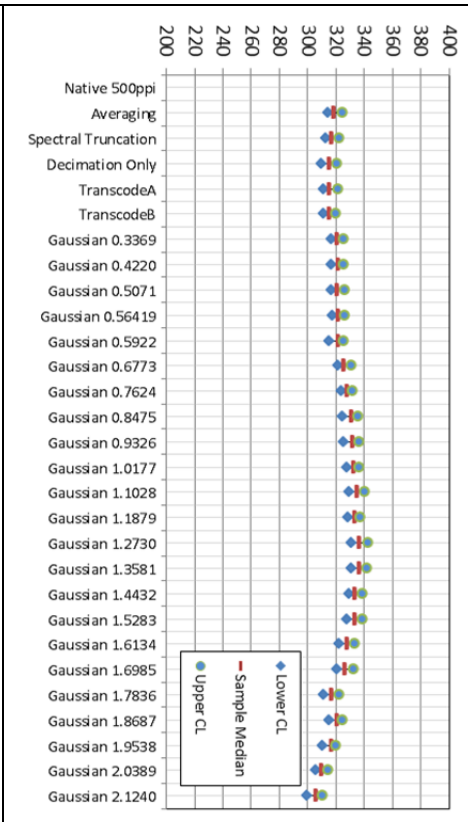
Figure 15 below shows 95 % Confidence Intervals for Automated Matching assessment based on 2000 replicates of mean rank, sample mean and sample median. Due to the operational nature of the research matcher used, native 500 ppi control images yielded very high match scores outside of the range of the other treatments. To improve readability of the charts, sample mean and sample median data for the Native 500 ppi Case have been omitted from the graphs. The 95 % confidence intervals for the sample mean were (Lower=714.5, Mean=721.1, Upper=727.9) and for the sample median CIs were (Lower=704, Median=712, Upper=721).



Mean Rank (Range 1-29, Higher is Better)



Sample Mean (Range 0 to 65535, Higher is Better)

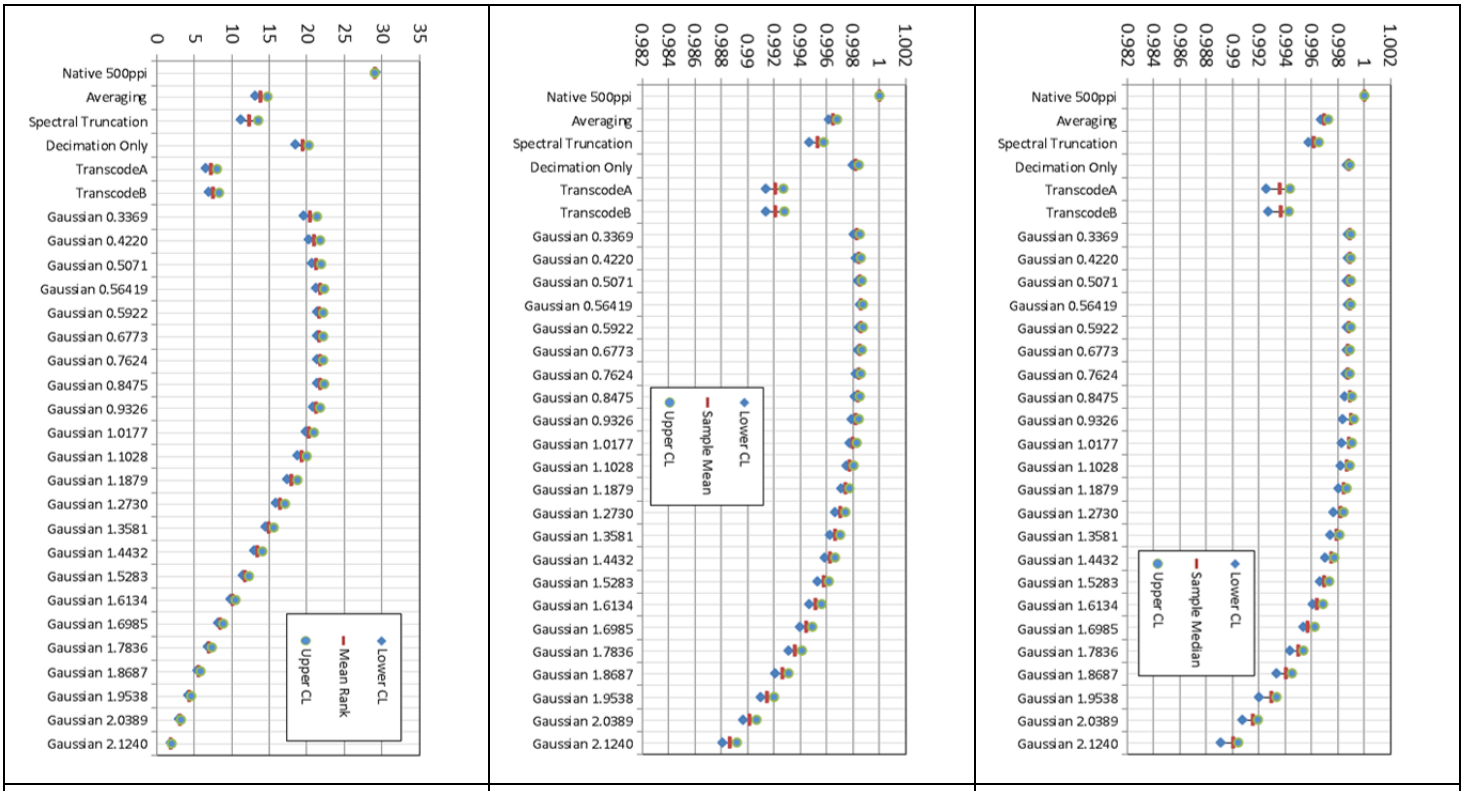


Sample Median (Range 0 to 65535, Higher is Better)

Figure 15 - Automated Matching Assessment 95 % Confidence Intervals for Mean Rank, Mean and Median using Bootstrap Analysis

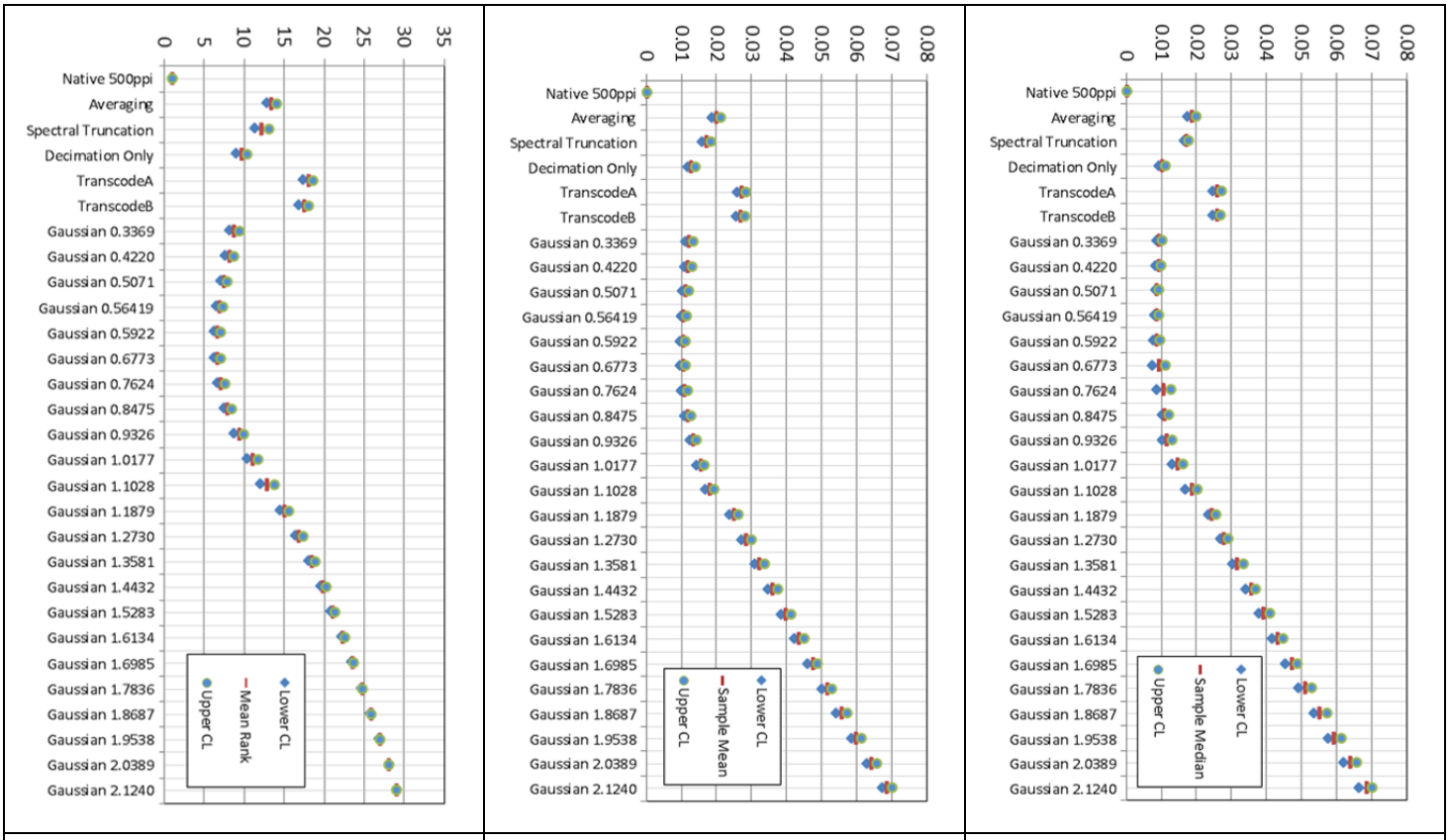
5.6.3. Signal Analysis and Assessment

Figure 16 and Figure 17 below shows 95 % Confidence Intervals for Frequency Spectrum Correlation and Root Mean Square Difference assessments based on 2000 replicates of mean rank, sample mean and sample median for each of these metrics.



Mean Rank (Range 1-29, Higher is Better)	Sample Mean (Range 0 to 1, Higher is Better)	Sample Median (Range 0 to 1, Higher is Better)

Figure 16 - Frequency Spectrum Correlation Analysis 95 % Confidence Intervals for Mean Rank, Mean and Median using Bootstrap Analysis



Mean Rank (Range 1-29)	Sample Mean (Range 0 to ∞)	Sample Median (Range 0 to ∞)

Figure 17 - Root Mean Square Difference Analysis 95 % Confidence Intervals for Mean Rank, Mean and Median using Bootstrap Analysis

5.6.4. Normalized Score Comparative Analysis

In another approach for the analysis of results, a method of score normalization was utilized. In this method, the scores from all metrics were normalized to fit into a scale of 0 (worst/lowest) to 1 (best/highest) and the resulting values were plotted against each other as show in. This analysis shows that RMSD is perhaps the most sensitive measure and exhibits the most fluctuation relative to the various treatment processes as it reacts to subtle changes to the image. Correlation analysis (frequency based) appears to exhibit the least amount of fluctuation. Examiner subjective analysis appears to have an optimal peak in the Gaussian range of (0.3369) and (1.1028).

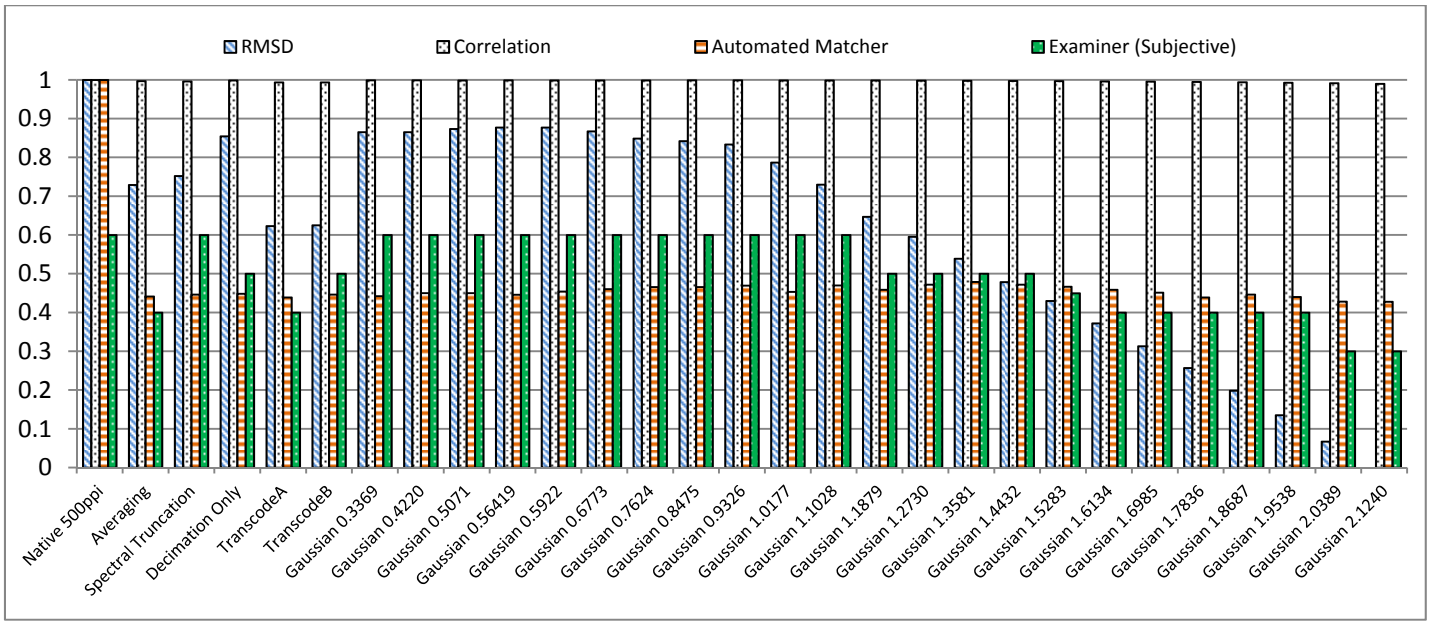


Figure 18 - Normalized Score Comparative Analysis

5.7. Other Considerations

In addition to the metrics explored in this study, there are other considerations that are significant when selecting a method. For example, each treatment may pose a significantly different impact from an operational perspective on certain aspects of the processing pathway through differences in processing needs and memory utilization as compared to other treatments and some treatments may have particular advantages over others in a given scenario. A comprehensive investigation of these operational ramifications was not conducted in the scope of this study but this section describes some of the notable differences.

The process of downsampling in the scope of this study involves three basic steps:

1. Decoding of compressed image data from a JPEG-2000-encoded file
2. Processing the image data in order to reduce the geometric dimensions (width and height) of the image by half
3. Encoding the resulting processed image data into a WSQ file

All of the treatments explored in this study follow those basic steps, though there is much variation among the processing of the image data in step 2 above. Some treatments even include two separate processing steps, such as the Gaussian treatments which involve the application of a processing filter followed by the decimation of the processed result in order to reduce the dimensions of the image.

Table 34 summarizes the fundamental differences between the treatments in terms of the above operational steps. As shown, Transcoded A and B require the fewest operational steps since only limited decompression of the 1000 ppi image is required, as well as no intermediate file saving of the decompressed image data, which may give it an advantage in terms of memory utilization as well as computational cycles (processing time). Gaussian on the other end of the scale requires the most operational steps, with an additional step required for decimation after filtering, but generally produces images that earned higher subjective ratings by expert examiners. It should also be noted that the Gaussian filter incurs a growing penalty as the kernel size increases. In the scope of this study the lower end of the Gaussian kernel was bound between (r=3) and (r=7), depending on the treatment.

Empirical testing of the various treatments (Figure 19 and Figure 20) confirms these results and shows that the Transcoded methods provide far superior throughput performance than all other methods examined in this study.

Table 34 - Generalized Comparison of Resource Requirements by Treatment

Treatment Class	Decompression	Decompressed File Save	Processing Filter	Post-Filter Decimation	Encoding to WSQ
Averaging	Full Required	Required	Required	Unnecessary	Required
Decimation	Full Required	Required	Unnecessary	Unnecessary	Required
Spectral Truncation	Full Required	Required	Required	Unnecessary	Required
Transcoded A	Partial Only	Required	Unnecessary	Unnecessary	Required
Transcoded B	Partial Only	Unnecessary	Unnecessary	Unnecessary	Required
Gaussian	Full Required	Required	Required	Required	Required

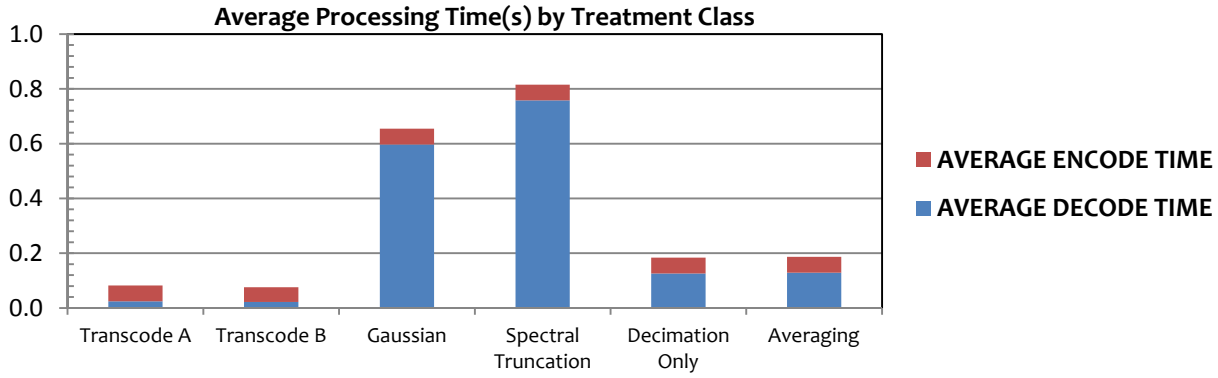


Figure 19 - Average Processing Time(s) by Treatment Class

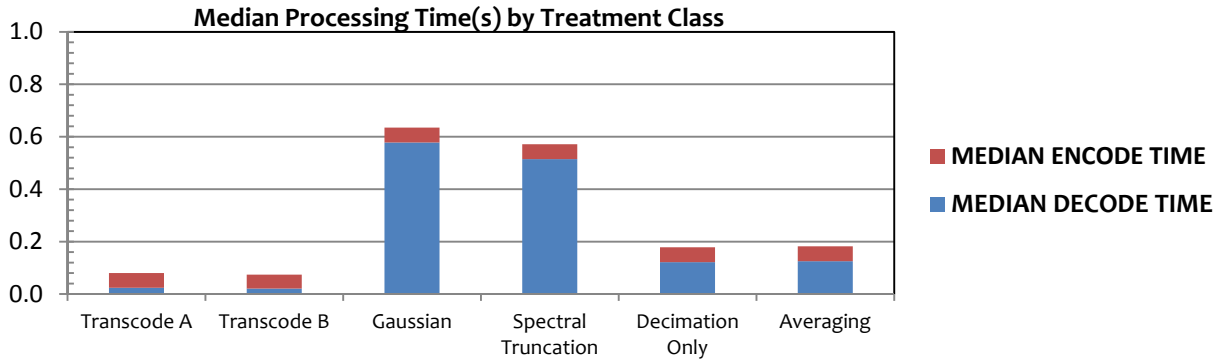


Figure 20 - Median Processing Time(s) by Treatment Class

6. Conclusions

Several different analysis methodologies/algorithms were used to identify an optimal strategy. These methods included various different fidelity measures (such as expert examiners or machine matcher score assessment) for each treatment as well as several different methods of fusing the results of these measures such as clustering ranks. As shown in Table 35, all the methods used to fuse the experimental measures indicated that the native 500 ppi control case yielded the best results versus a 1000 ppi image that has been downsampled to 500ppi. The next best method identified by each fusion methodology was one of the Gaussian treatments. The results from Table 35 also reinforce anecdotal evidence that it is best to capture at the resolution one intends to operate in rather than post-process the image further where the native 500 ppi control case bested images generated by downsampling treatments.

Table 35 - Summary of Findings

Measure Fusion Method	Best Treatment Case (Optimal)	Second Best Treatment Case
Clustered Rank Summation (Optimistic Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.6773, r = 3
Clustered Rank Summation (Conservative Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.9326, r = 4
Ordinal Rank Summation (Optimistic Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.8475, r = 4
Ordinal Rank Summation (Conservative Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.9326, r = 4
Mean Rank Summation (Optimistic Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.8475, r = 4
Mean Rank Summation(Conservative Decimation)	Native 500 ppi Control Case	Gaussian σ = 0.7624, r = 3
Expert Examiner Winner-Take-All	Native 500 ppi Control Case	Gaussian σ = 0.8475, r = 4

Based on the findings in this study, the Gaussian treatment (σ = 0.8475, r = 4) is identified as an optimal all-around downsampling strategy for the measure fusion methods of Ordinal Rank Summation-Optimistic, Mean Rank Summation-Optimistic and Expert Examiner Winner-Take-All (which further fuses optimistic and conservative decimation strategies). Where Gaussian treatment (σ = 0.8475, r = 4) is not identified as optimal following the control case, it is identified as third best (or tied as such) counting the control case as the best method and is not a radical compromise in terms of the fidelity measures that didn't identify it as the top candidate.

Looking deeper into the various individual measures did uncover some surprises. In the case of expert examiner quality assessment, the downsampled images yielded better subjective quality than the native 500 ppi control case which indicates that some signal noise reduction may ultimately impact the perception of image quality by expert examiners.

With respect to objective examiner assessments of the images where the examiner was asked to make an identification decision between a downsampled image and the native 500 ppi version of that image, a statistically significant difference could not be established. This reinforces findings from NISTIR-7778 [NISTIR7778] where the examiners could nearly universally detect degradation in the various aspects of the fingerprint ridge structure as a result of lossy compression, but this effect did not appear to measurably impact their ability to conduct their identification tasks. This is analogous to operating a motor vehicle with a "dirty windshield", where the driver's ability and experience mitigates any effect from occlusions large or small in their field of view.

An examination of automated matching shows that the Control Case outperformed all other treatments, followed by several Gaussian treatments including (σ = 1.2730, r = 5, Odd-Even) for the optimistic decimation strategies and (σ = 1.1879, r = 5, Even-Odd) for the conservative decimation strategy. Based on the results of this study however, it appears that a strategy that optimized performance of the research matcher used for this study (via very large Gaussian filters) yielded degradation from subjective assessment of the image by expert examiners. While the images favored by the matcher were not the worst in terms of expert examiner assessments, they did rank far lower by examiners. While it is possible to apply further filtering to an image, it is not possible undo signal loss as a result of Gaussian filtering therefore a potential strategy to favor both the expert examiners as well as the matcher may involve applying one pass of Gaussian filtering in the downsampling process, and a second pass just prior to identifying feature vectors on the fingerprint images.

Examination of system throughput however points to Transcoded A and Transcoded B as the optimal solution in terms of operational efficiency and lower resource demands by eliminating several processing steps in the downsampling pathway. The elimination of processing steps by operating on the wavelet data directly helped avoid some of the intermediate steps needed for traditional downsampling and avoid the associated throughput penalty for these intermediate steps. Cursory studies show that these methods may be up to 10 times faster than the more computationally intensive Gaussian methods.

Finally, with regard to Gaussian filter sizes, it was shown that the automated matching favors Gaussian cases with larger sigma & radius values while other fidelity measures favor smaller sigma & radius value indicating that cases that are optimized for automated matching specifically may not be optimal for other fidelity measures including expert human perception.

7. Discussions and Future Work

This study showed that a certain amount of noise reduction in the image may have a positive effect both on the subjective analysis of the image, as well as automated machine matcher efficacy. Further research is needed to see if the introduction of controlled/targeted high frequency noise filtering yields a positive net benefit in the processing and operational pathway of friction ridge imagery, especially in terms of impact on expert examiner identification error rates.

Another finding from the phase-one decimation portion of analysis used in several segments of this study showed that the scanning process may be more deterministic than previously thought. Specifically, certain post-processing or conditional operations may not exhibit random behavior as previously thought (given expected outcome probabilities for the number of cases). Does the scanner introduce a bias into the sampling of fingerprints that favors one decimation method over the others? Is the effect limited to the scanner used in the present study or do other scanners exhibit similar effects? In addition, preliminary experiments do indicate limits of scanner repeatability that have not been quantified and may pose certification issues, particularly among aging devices. Accordingly, further study of factors impacting the scanning process is planned.

References

Publications and Reports

BERGEN	Bergen, J. R., Anandan, P., Hanna, K. J., and Hingorani, R. (1992) "Heirarchical model-based motion estimation" in Proceedings of Second European Conference on Computer Vision, Springer-Verlag, p. 237-252.
CHAMBERS	Chambers, John; William Cleveland, Beat Kleiner, and Paul Tukey (1983). Graphical Methods for Data Analysis. Wadsworth
CIA	"The World Factbook", https://www.cia.gov/library/publications/the-world-factbook/geos/us.html, Retrieved 2010-12-06.
CONOVER	"Practical Nonparametric Statistics", Third Edition, Wiley, 1999, pp. 367-373.
CURRAN-EVERETT	Curran-Everett, Douglas (2000). "Multiple comparisons: philosophies and illustrations", American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 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.
FITZPATRICK	Fitzpatrick, M. et al. 1994, "WSQ Compression / Decompression Algorithm Test Report", IAI Annual Conference.
HOLLANDER	Hollander, M., Wolfe, D. A. (1999). Non-Parametric Statistical Methods, 2nd Ed, John Wiley & Sons:NY, 787 pages.
LIBERT	"A 1D Spectral Image Validation/Verification Metric for Fingerprints". Libert, J.M.; Grantham, J.; Orandi, S. August 19, 2009. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=903078. Retrieved 2011-01-12.
LIKERT	Likert, R. (1932). *A Technique for the Measurement of Attitudes,* Archives of Psychology 140, 55.
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.
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.
NISTIR7778	Orandi, S., Libert, J. M., Grantham, J. D., Ko, K., Wood, S.S., Wu, J. Effects of JPEG2000 Image Compression on 1000 ppi Fingerprint Imagery, NIST Interagency Report 7778, National Institutes of Standards and Technology, Gaithersburg, MD. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=911122. Retrieved from on 09/01/2012.
SOKAL	Sokal, Robert R. and Rolf, F. James (1969) Biometry: The Principals and Practice of Statistics in Biological Research, Third Edition, W. H. Freeman & Co., NY, 887 pages.
WILCOXON	Wilcoxon, Frank (Dec 1945). "Individual comparisons by ranking methods". Biometrics Bulletin 1 (6): 80-83.
WRIGHT	Wright, Paul, S. P-values for simultaneous inference. Biometrics, 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, Journal of Research of the National Institute of Standards and Technology, 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

ANSI/NIST-ITL 1-2011	ANSI/NIST-ITL 1-2011, NIST Special Publication 500-290 Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information. http://www.nist.gov/itl/iad/ig/ansi_standard.cfm . Retrieved 2012-09-11.
ISO/IEC 15948:2004	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.
ISO/IEC 19794-1:2011	ISO/IEC 19794-1:2011- Information technology -- Biometric data interchange formats -- Part 1: Framework. http://www.iso.org/iso/catalogue_detail.htm?csnumber=50862. Retrieved 2012-09-10.
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 -- JPEG2000 image coding system: Core coding system". http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=27687. Retrieved 2009-11-01.
WSQ	"WSQ Gray-Scale Fingerprint Image Compression Specification" Version 3.1. https://www.fbibiospecs.org/docs/WSQ_Gray-scale_Specification_Version r = 3, 1.pdf. Retrieved 2010-01-11.

Appendix A. - **Dataset 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 both 500 ppi and 1000 ppi by NIST personnel under controlled conditions and without repositioning the cards between scans at 500 ppi and 1000 ppi. 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 are introduced into the original set.

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 of each subsample was to present a sufficiently large set of each subsample so that they avoid the potential statistical bias of very small subsamples. That is, all subsamples are 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 rolled images. Each subject contributed exactly one fingerprint to the dataset.

Table 36 - Ink Card Scan Data classification by Impression Type

All Data				
Impression Type	Males	Females	Right	Left
Rolled Single Finger	183	17	100	100

Table 37 - Ink Card Scan Pattern Classification for Single Finger Images by Impression Type

Data From Females (Single Finger)				Data From Males (Single Finger)			
Pattern Class	Rolled	Right	Left	**Pattern Class**	Rolled	Right	Left
Arch	1	0	1	**Arch**	16	9	7
Loop	8	4	4	**Loop**	104	53	51
Whorl	8	5	3	**Whorl**	63	29	34
Total	17	9	8	**Total**	183	91	92

Table 38 - 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	0	0	0	0	0	0	1	0	0	0	1
Loop	1	0	2	1	0	0	3	0	0	1	8
Whorl	2	2	1	0	0	0	1	0	1	1	8
Total	3	2	3	1	0	0	4	0	1	2	17

Table 39 - 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	1	6	1	1	0	2	1	3	0	1	16
Loop	10	8	10	8	17	9	4	12	10	16	104
Whorl	6	4	6	10	3	9	10	5	9	1	63
Total	17	18	17	19	20	20	15	20	19	18	183

Table 40 - Gender Breakdown for Data

Subjects by Gender and Race							
	Males	**Females**	**White**	**Black**	**Hispanic**	**Asian**	**Native American**
Ink Capture Dataset	183	17	97	89	7	6	1

Table 41 - Age Breakdown for Data

Subjects by Age									
	Under 18	**18-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50+**	**UNKNOWN**
Ink Card Scan Dataset	7	128	31	15	5	4	5	2	3

Table 42 - Other Metadata: Height and Weight

Subjects by Height and Weight											
	Height					**Weight**					
	<5'0"	**5'0" - 5'5"**	**5'6" - 5'11"**	**6'0"+**	**UNKNOWN**	**<100 lbs**	**100 - 149 lbs**	**150 - 199 lbs**	**200 - 249 lbs**	**250+ lbs**	**UNKNOWN**
Ink Card Scan Dataset	0	20	127	52	1	0	58	115	21	5	1

Table 43 - Other Metadata: Eye Color

Subjects by Eye Color						
	Brown	**Black**	**Blue**	**Green**	**Hazel**	**UNKNOWN**
Ink Card Scan Dataset	147	3	30	8	11	1

Appendix B. - **Equipment Used for Study**

The equipment utilized for the processing of the image data described in this study 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**

A different set of equipment was used in the data collection portion of this study and consisted of the following:

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

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

3 x 24 inch Monitors: The IPS-panel monitors are connected via DVI-D and calibrated (see below) for optimal accuracy and consistency. The monitors were operated at their native resolution of 1920x1200, yielding a dpi measurement of approximately 94.3ppi. The zoom functionality in software allowed the examiners to zoom in and out of the image, and view them in the range of approximately 10x to 50x given these specific monitors.

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

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

Appendix C. - **Gaussian filter code**

```
// The Following code segment has been developed in the C#.NET platform

// Runs image through a Gaussian filter with the parameters specified
        public void GaussianFilter(double sigma, int radius)
        {
            Console.WriteLine("\tRunning Gaussian filter...");

            // Construct Gaussian function

            int length = radius * 2 + 1;
            double[,] gaussian = new double[length, length];
            Point center = new Point(radius, radius);

            // set distances from center

            for (int x = 0; x < length; x++)
                for (int y = 0; y < length; y++)
                    gaussian[x, y] = Distance(new Point(x, y), center);

            // run distances through Gaussian function
            double sum = 0;
            for (int x = 0; x < length; x++)
                for (int y = 0; y < length; y++)
                {
                    gaussian[x, y] = Gaussian(gaussian[x, y], 0, sigma);
                    sum += gaussian[x, y];
                }

            for (int x = 0; x < length; x++)
                for (int y = 0; y < length; y++)
                    gaussian[x, y] = gaussian[x, y] / sum;

            // Apply filter to image
            int height = bitmap.Height;
            int width = bitmap.Width;
            Bitmap tmpbmp = new Bitmap(width, height);
            for (int y = 0; y < height; y++)
                for (int x = 0; x < width; x++)
                {
                    // get surrounding pixels
                    int[,] area = new int[length, length];
                    for (int y2 = 0; y2 < length; y2++)
                        for (int x2 = 0; x2 < length; x2++)
                        {
                            int xbound = x - radius + x2;
                            int ybound = y - radius + y2;
                            if (xbound >= 0 && xbound < width && ybound >= 0 && ybound < height)
                                area[x2, y2] = array[x - radius + x2, y - radius + y2];
                            else // if filter goes off the edge, use nearest pixel
                                area[x2, y2] = array[x, y];
                        }

                    // find weighted sum
                    double wsum = 0;
                    for (int x2 = 0; x2 < length; x2++)
                        for (int y2 = 0; y2 < length; y2++)
                            wsum += area[x2, y2] * gaussian[x2, y2];
                    tmpbmp.SetPixel(x, y, Value2Color((int)(wsum)));
                }

            // save
            bitmap = tmpbmp;
            InitArray();
        }

        // Downsamples image by 50% in both dimensions
        public void Downsample50(int dstype)
        {
            Console.WriteLine("\tDownsampling...");

            int height = bitmap.Height;
            int width = bitmap.Width;
            Bitmap dsbitmap;
            switch (dstype)
            {
                case Constants.EE: // save even columns and rows
```

```
            dsbitmap = new Bitmap((bitmap.Width + 1) / 2, (bitmap.Height + 1) / 2);
            for (int y = 0; y < height; y++)
                if (y % 2 == 0)
                    for (int x = 0; x < width; x++)
                        if (x % 2 == 0)
                            dsbitmap.SetPixel((x + 1) / 2, (y + 1) / 2, Value2Color(array[x, y]));
            break;
        case Constants.EO: // save even columns and odd rows
            dsbitmap = new Bitmap((bitmap.Width + 1) / 2, bitmap.Height / 2);
            for (int y = 0; y < height; y++)
                if (y % 2 == 1)
                    for (int x = 0; x < width; x++)
                        if (x % 2 == 0)
                            dsbitmap.SetPixel((x + 1) / 2, y / 2, Value2Color(array[x, y]));
            break;
        case Constants.OE: // save odd columns and even rows
            dsbitmap = new Bitmap(bitmap.Width / 2, (bitmap.Height + 1) / 2);
            for (int y = 0; y < height; y++)
                if (y % 2 == 0)
                    for (int x = 0; x < width; x++)
                        if (x % 2 == 1)
                            dsbitmap.SetPixel(x / 2, (y + 1) / 2, Value2Color(array[x, y]));
            break;
        case Constants.OO: // save odd rows and columns
            dsbitmap = new Bitmap(bitmap.Width / 2, bitmap.Height / 2);
            for (int y = 0; y < height; y++)
                if (y % 2 == 1)
                    for (int x = 0; x < width; x++)
                        if (x % 2 == 1)
                            dsbitmap.SetPixel(x / 2, y / 2, Value2Color(array[x, y]));
            break;
        default:
            throw new Exception("Illegal value for downsampling type. Needs to be 1-4.");
    }

    // save
    bitmap = dsbitmap;
    InitArray();
}

private void InitArray()
{
    array = new int[bitmap.Width, bitmap.Height];
    for (int x = 0; x < bitmap.Width; x++)
        for (int y = 0; y < bitmap.Height; y++)
            array[x, y] = bitmap.GetPixel(x, y).R;
}

// converts int value to a Color
private Color Value2Color(int pixval)
{
    int ARGB_val;
    Color color;

    ARGB_val = pixval * 256 * 256 + pixval * 256 + pixval;
    color = Color.FromArgb(ARGB_val);

    return color;
}

private double Distance(Point a, Point b)
{
    return Math.Sqrt((a.X - b.X) * (a.X - b.X) + (a.Y - b.Y) * (a.Y - b.Y));
}

// returns value of Gaussian function with given x, mu, and sigma
private double Gaussian(double x, int mu, double sigma)
{
    return (1 / Math.Sqrt(2 * Math.PI * sigma * sigma)) * Math.Exp(-((x - mu) * (x - mu)) / (2 * sigma * sigma));
}
```