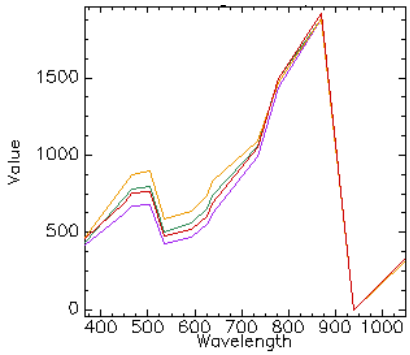
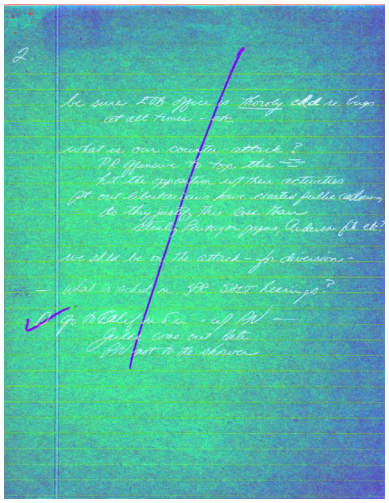
LIBRARY OF CONGRESS

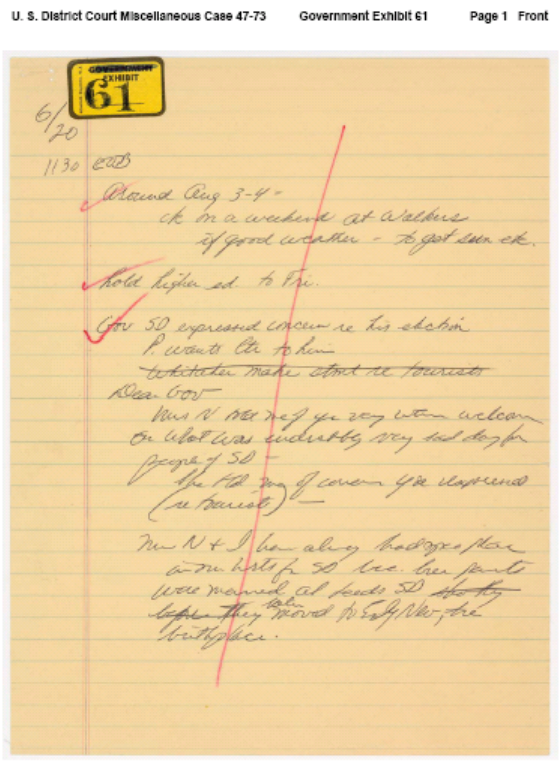
PRESERVATION DIRECTORATE
PRESERVATION RESEARCH AND TESTING DIVISION
101 INDEPENDENCE AVENUE, SE
WASHINGTON, DC 20540-4560

HYPERSPECTRAL EXAMINATION OF H.R. HALDEMAN PAPERS 2009-2011

Fenella G. France PhD, MBA, FAIC







6/7/2011

Contents

Research Project: H. R. Haldeman Papers 6/20, 6/21 (1972)	
PRTD Scientist:	Fenella G. France, Library of Congress
Document Institution:	National Archives and Records Administration (NARA)
NARA POC:	Margaret Kelly
Other Collaborators:	Treasury Inspector General for Tax Administration (TIGTA) Bureau of Alcohol, Tobacco, Firearms and Explosives (BAFTE) (abbreviated ATF)

Executive Summary

- Research proved the usefulness of hyperspectral imaging for non-invasive analysis of the Haldeman papers (6/20, 6/21, 1972). This served to answer a range of questions and issues that had been raised as part of the investigation.

- Overlays of indented ESDA and HSI images provided some success at increasing the contrast to enhance readability of indented writing observed after the ESDA examination. White-light interferometric optical profilers could potentially be applied to hyperspectral images to enhance the readability of indented writing.

- Spectral curves obtained from the hyperspectral imaging indicated some differences between the text on the body of the page and the top corner margin page number and date text for the two pages from 6/20.
  No observable difference was apparent for the paper substrate used for the pages examined from 6/20 and 6/21.

- Additional processing using techniques such as PCA and pseudocolor processing enhanced and revealed the presence of finger prints on some of the pages examined.
  A sample of how the intensity of pressure in handwriting can be examined was included.

- Further processing that included masking regions to take the statistics from the background paper (with and without lines and other features) revealed previously unseen features including the transfer of the red “strike-through” line from page 2 to page 1 of 6/20, as well as another fingerprint not detected with other techniques.

- Full spectrum color images were created for all pages imaged by the hyperspectral imaging system to create high resolution images that closely resemble the color and features of the original document.

Project Objective

Forensic document analysis of Handwritten notes of H.R. Haldeman prepared during his 11:30am Executive Office Building meeting with President Nixon on June 20, 1972 that correspond to the audio tape bearing the "18 ½ Minute Gap" erasures.

Background

The forensic document examination occurred due to review of the handwritten notes of H.R. Haldeman, where it had been suggested that the two pages contained insufficient content to represent the entirety of the meeting and that additional pages bearing content corresponding to the audio tape erasures may have been removed prior to submission of the materials on subpoena Examination of the documents was requested in order to render legible any indented writing that may be present, particularly on the second page.

NARA convened a panel of experts from Federal laboratories at BAFTE, TIGTA and the Library of Congress to conduct an onsite review of the questioned documents and associated contemporaneous Haldeman materials, to therefore jointly concur and determine subsequent examinations of the relevant documents. This initial review occurred at NARA – College Park, MD 9:00am – 12:00noon 29 October 2009. It was agreed by the panel to begin the examinations at the Library of Congress with the most non-invasive technique, hyperspectral imaging, before continuing to techniques that were potentially more invasive.

Seven pages were selected for this initial hyperspectral imaging examination in order to achieve four goals:

1. Reveal any latent impressions on the pages
2. Compare ink entries on the selected pages
3. Compare paper substrates of the selected materials
4. Utilize the unique spectral response of materials (both substrate – paper, and media – inks) to capture any additional information.

Documents under examination included the following:
Government Exhibit 61 U.S. District court Miscellaneous Case 47-73
6/20 "11^{30}" EOB (2 pages)

Richard Nixon Presidential Library / White House Special Files / Staff Member and Office Files / H.R. Haldeman / Haldeman Notes / H notes/ Box 45 / Folder 4.
April – June 1972 [May 9 1972 – June 30, 1972] Part II
6/21 "Be sure S is with Burns" [2 pages document withdrawn N5/ 173]
6/21 "13^{30} D econom grp Mon" [3 pages]

Research Procedure / Methodology

Hyperspectral Imaging System

The MegaVision-Equipoise imaging system was installed at the Library in 2008 and has been used extensively over the past two years to image collection items, providing the range of spectral data needed for Library collections research. Utilizing a 39 megapixel monochrome camera (7216 x 5412) E6 back, and APO-Digitar 5, 6/120 lens, the camera setup is integrated with light emitting diode (LED) illumination panels that cover the spectral range of 365nm to 1050nm – extending from the ultraviolet (UV), through the visible spectrum and into the near infrared (IR).

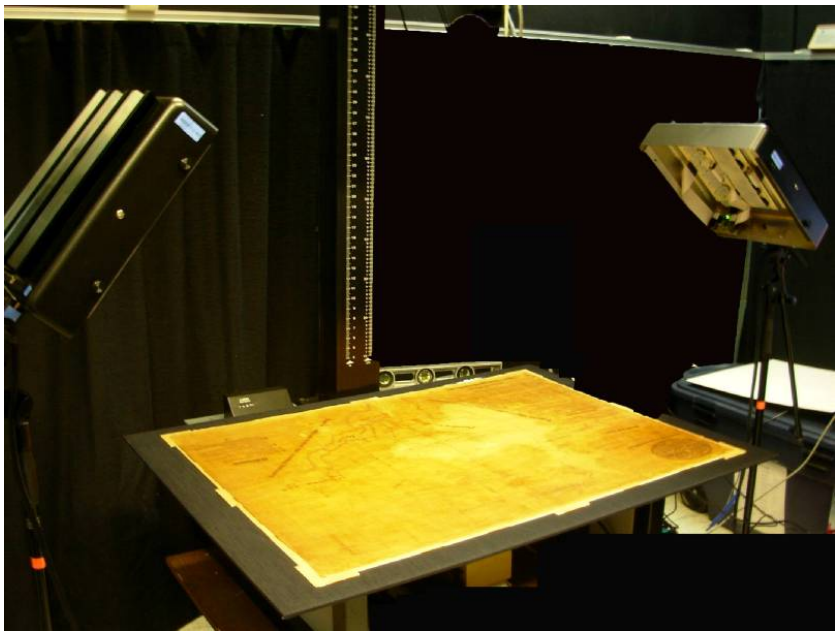


Figure 1. Hyperspectral imaging set-up: MegaVision 39 megapixel monochrome camera with integrated LED illumination panels

The image capture software has been customized to address the needs of information and metadata required for the cultural heritage field. The range of imaging modalities allows greater image analysis and downstream processing through capturing UV, visible and near IR spectral data in reflected, transmitted and raking (side-lighting) orientations. This approach ensures minimal handling of fragile items, with capture of extensive imaging data. All images are accurately registered, enabling almost unlimited combinations of spectral wavebands for processing and characterization.

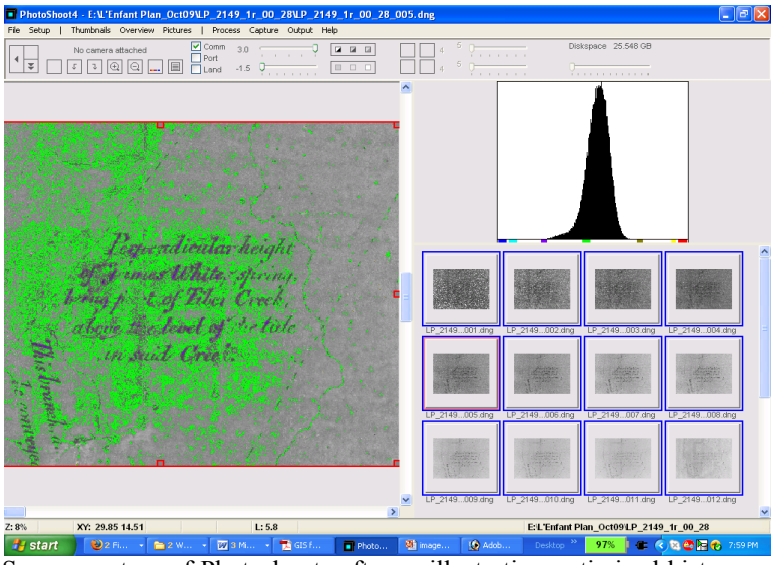


Figure 2. Screen capture of Photoshoot software illustrating optimized histograms for each spectral waveband and a selection of the images that form the data cube.

Hyperspectral imaging (HSI) involves taking a series of digital images in over thirteen different nanometer bandwidths of the visible and non-visible spectrums to detect variances in components or material composition of the object being imaged. The below table provides the specific wavelengths and light duration for each individual LED. This duration is optimized as shown in the above screen capture.

N-shot Table

Peak Wavelength (nm)	Color	Light Duration (sec)
365	UV	3.2
450	Royal Blue	1.0
465	Blue	1.0
505	Cyan	0.8
535	Green	0.7
592	Amber	0.9
625	Red-Orange	0.2
638	Red	0.17
730	IR	2.0
780	IR	2.0
850	IR	2.0
940	IR	15.0
1050	IR	30.0
470	Raking Blue Left	2.0
910	Raking IR Left	4.0
470	Raking Blue Right	2.0
910	Raking IR Right	4.0

Images were combined at either any one wavelength or combination of wavelengths. The resulting images were digitally combined, subtracted or processed to form images for scientific or scholarly analysis. These spectral images contained a wealth of information, but required significant interpretation to process and analyze the collected data. Advances in technology have greatly impacted the field of spectral imaging, which have been further enhanced through improvements in the process of image capture techniques and specific post-image processing. Open source and off-the-shelf packages including ENVI-ITT, Image J and Photoshop CS3, and CS5 software have been utilized in the processing to enhance obscured and hidden features.

All imaging metadata was captured and attached to each raw and successive processed image files. The raw image file format was generated as an open source standard digital negative graphic (DNG) and after a "flattening" process to correct for any minor variation in light intensity across the imaging surface, a 16-bit tagged image file format (TIFF) image cube is generated for analysis and further processing. Data were stored using broadly accepted standards for later retrieval and analysis with this focus on accurate, standardized metadata and file formats[7]. The system provides the foundation for layering levels of spectral data. This creates a composite cube of data from non-invasive imaging that can map and link additional layers of information with other analytical modalities. The control of potentially damaging light levels and fluctuating environmental factors found in other analytical techniques enables this spectral imaging technique to minimize the impact on fragile collection items. Integration with spectral reference libraries and advanced image processing software facilitates identification of chemical composition, changes due to deterioration to reveal areas at risk, changes to the document from previously undetected historic details, detection of change and impact from environmental conditions and document archaeology.

Image Acquisition

All pages were imaged at a resolution of 479dpp (dots per pixel). Imaging took place on 11/19/09 from 9:00am until approximately 3:30pm. Files were named and organized with the following structure:

File structure key:
Example 620_1130_P1_1r_00
620: 6/20 (date)
1130: time or other identifier at top of page
P1: page number in notes
1r: recto (front) or verso (back) of page

The full list of imaged pages and file names is as follows:
620_1130_P1_1r_00
620_1130_P1_1v_00
620_1130_P2_1r_00
620_1130_P2_1v_00

621_besure_P1_1r_00

621_besure_P1_1v_00
621_besure_P2_1r_00
621_besure_P2_1v_00

621_1330_P1_1r_00
621_1330_P1_1v_00
621_1330_P2_1r_00
621_1330_P2_1v_00
621_1330_P3_1r_00
621_1330_P3_1v_00

As noted above the raw files were captured as DNGs and then underwent "flat fielding". This involves imaging a white blank reference sheet to assess the light distribution across the surface of the document, and the flat field file is then subtracted from the original to remove any minor variations in light distribution across the surface of the document being imaged. These "flattened" files were captured as 16-bit TIFF images.

All pages were imaged in two orientations – initially with the long edge of the page oriented North-South to the camera (lighting being east-west). Each document was then rotated 90° clockwise to assess the capabilities of raking (side-lighting) at both orientations to detect any topographical variations / indentation on the pages. The angle of lighting for the LED illumination panels for the N-shot table was 45°. Side lighting (raking illumination) was at a 12° angle. As in figure 3, all pages were imaged with a scale rule and Gretag Macbeth color checker. This allowed the creation of full spectrum color images.

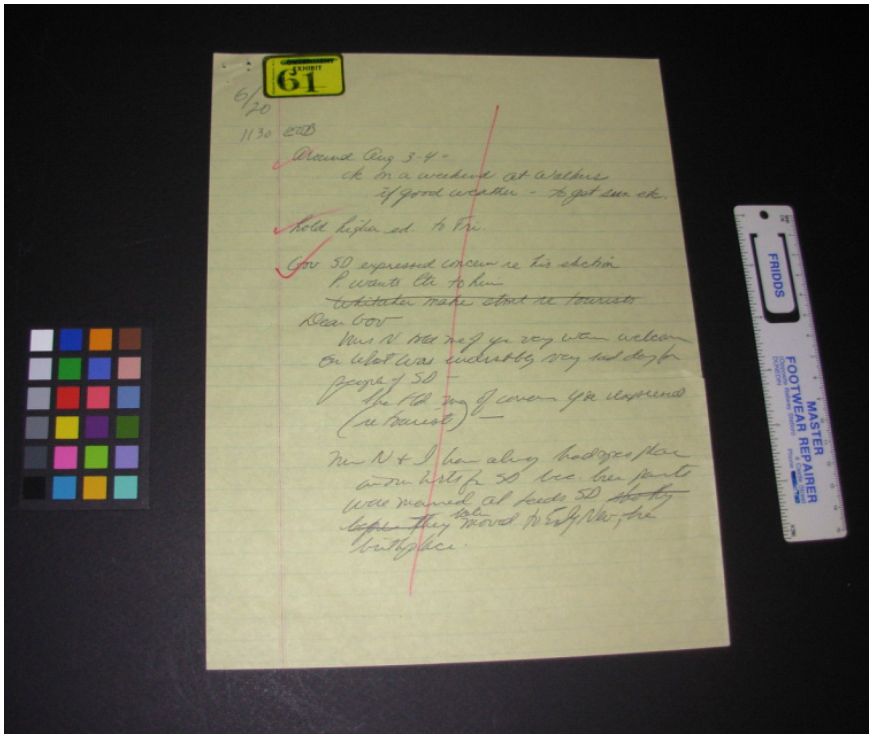


Figure 3. Imaging layout for Haldeman papers

Results

Latent or Indented Images

Examinations were performed by TIGTA and BATFE forensic examiners using Electrostatic Detection Apparatus (ESDA) on both 12/11/09 and 01/13/2010. Analyses undertaken on the page entitled 6/20 1130 EOB Page 2 under the conditions of Front Humidified, TIGTA beads + toner revealed the presence of some additional indented writing. These ESDA generated images were compared with hyperspectral images and underwent image processing to attempt to increase the contrast and enhance any features located through this technique.

The below image illustrates the overlay of the two image files with other processing variations to enhance the indented text on following pages.

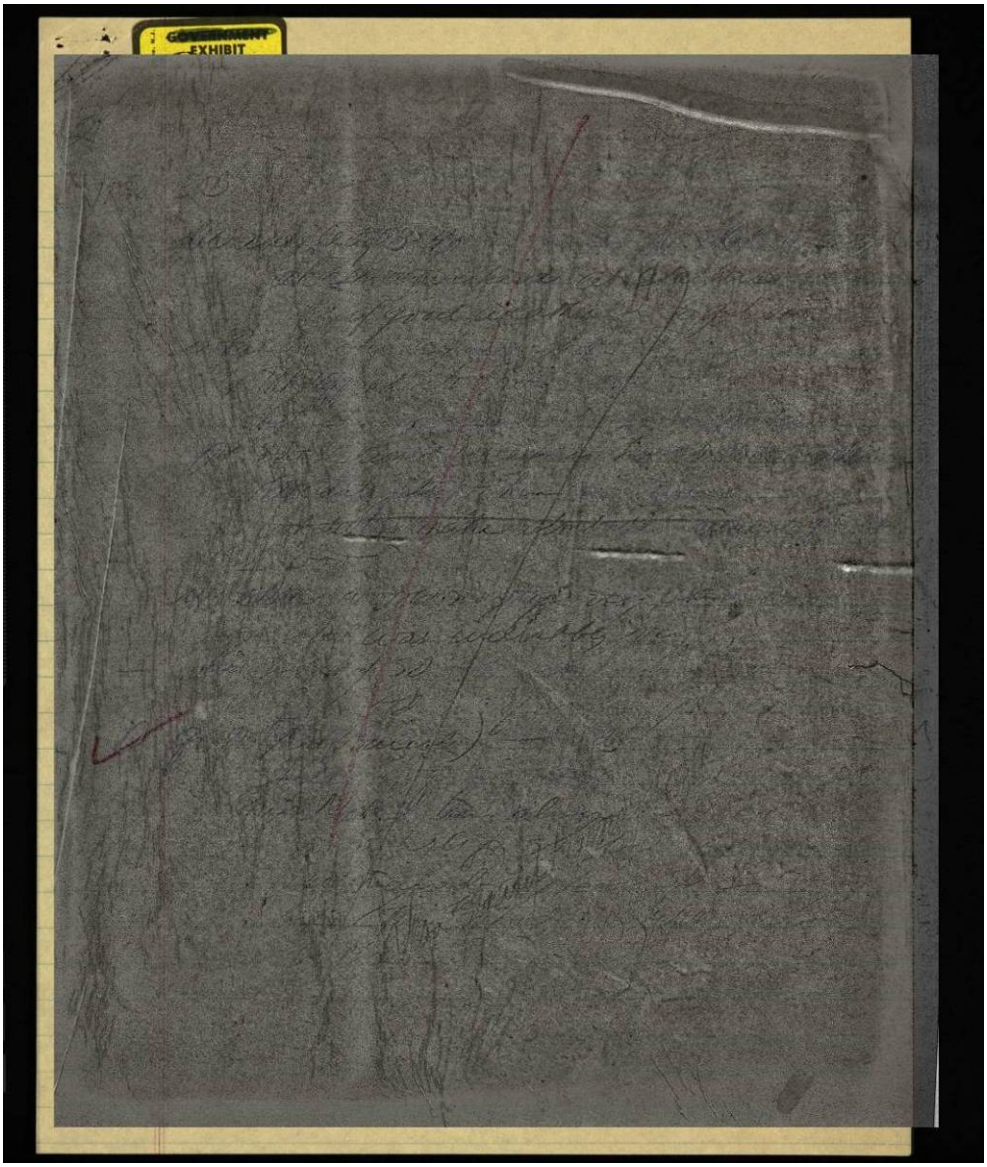


Figure 4. Overlay of Page 2 ESDA Image with Page 1 HSI Full Spectrum Images

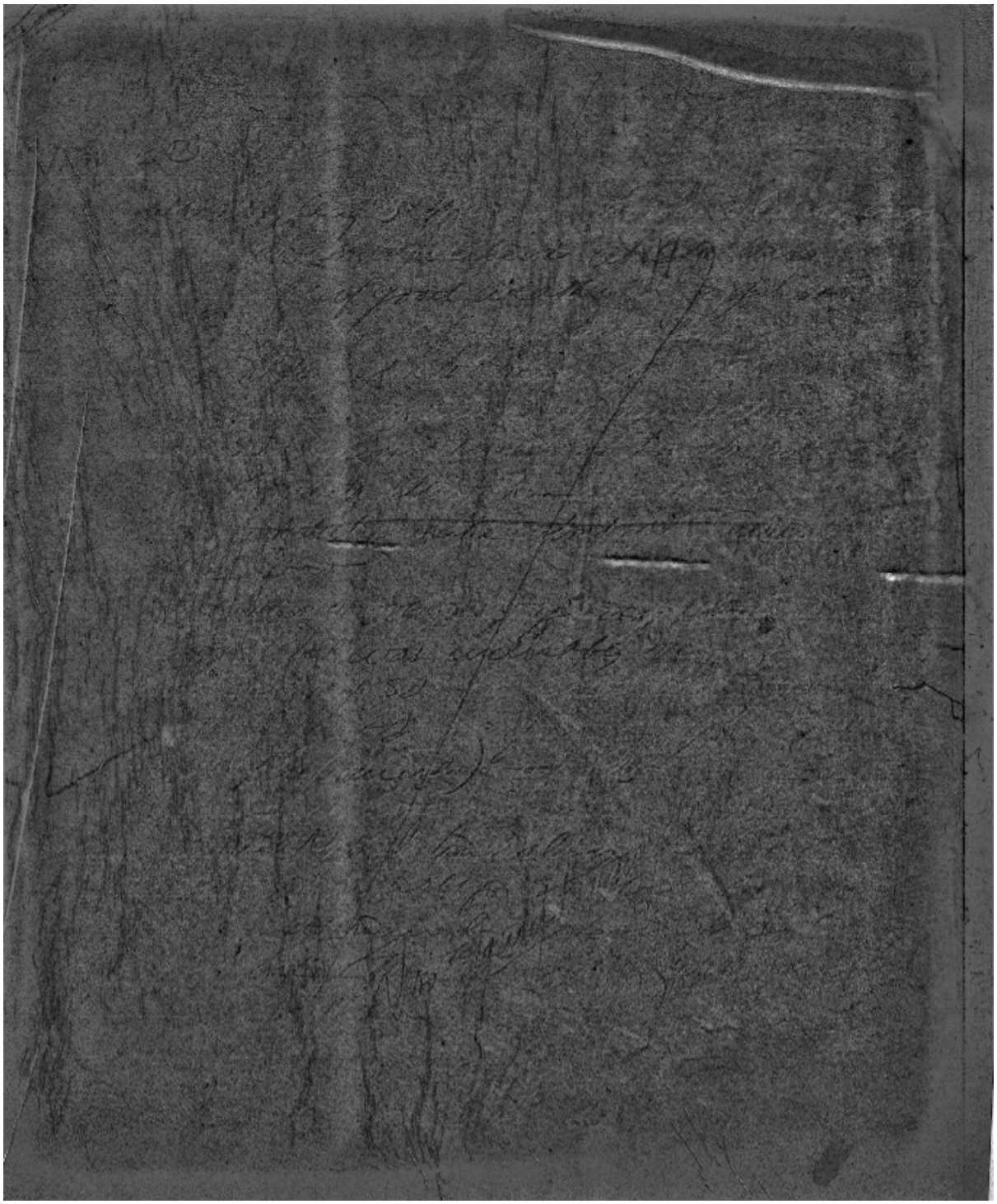


Figure 5. Close-up of Page 2 Lower Region with Indented Writing

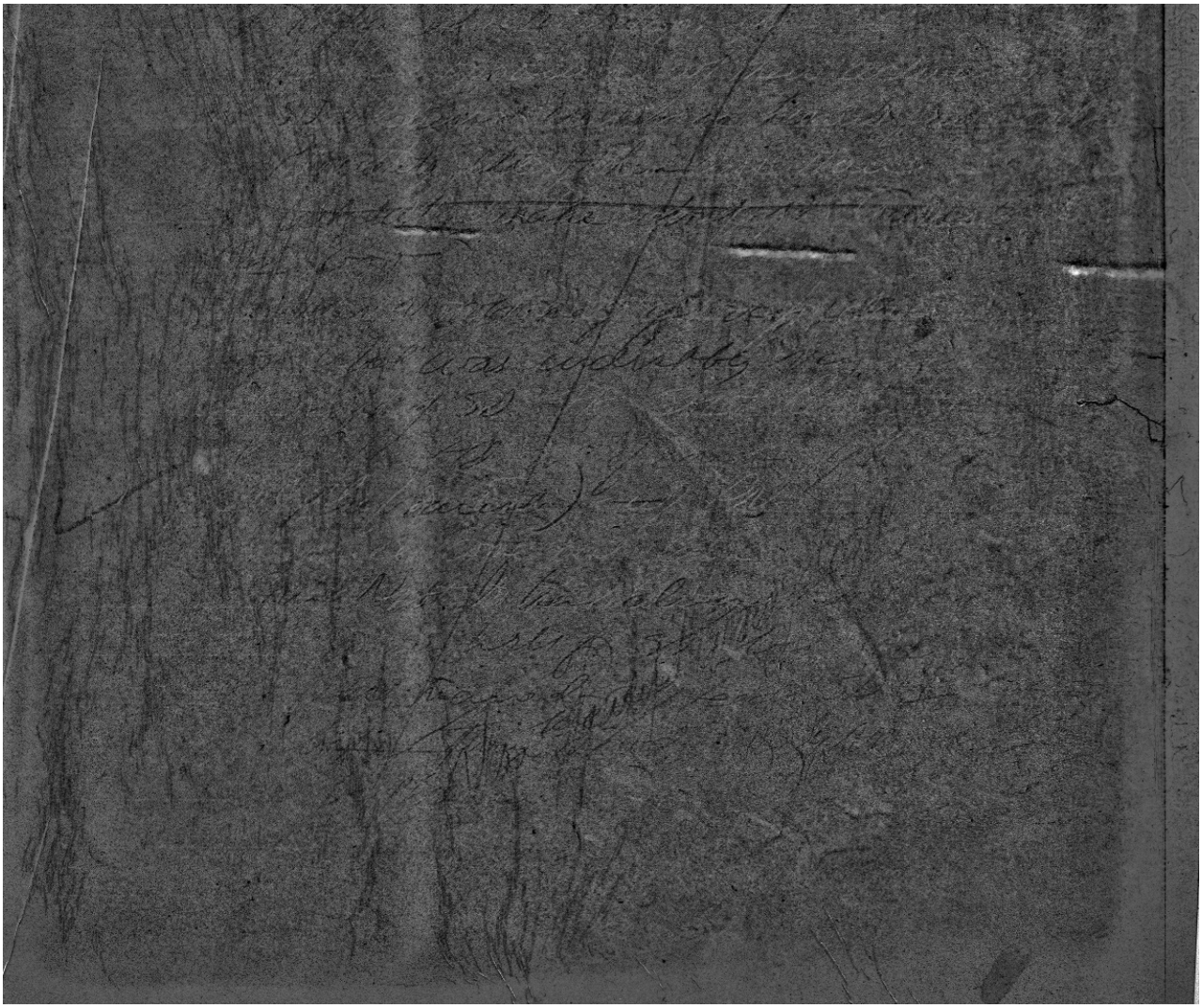


Figure 6. Close-up of Page 2 Lower Region with Indented Writing - Exclusion

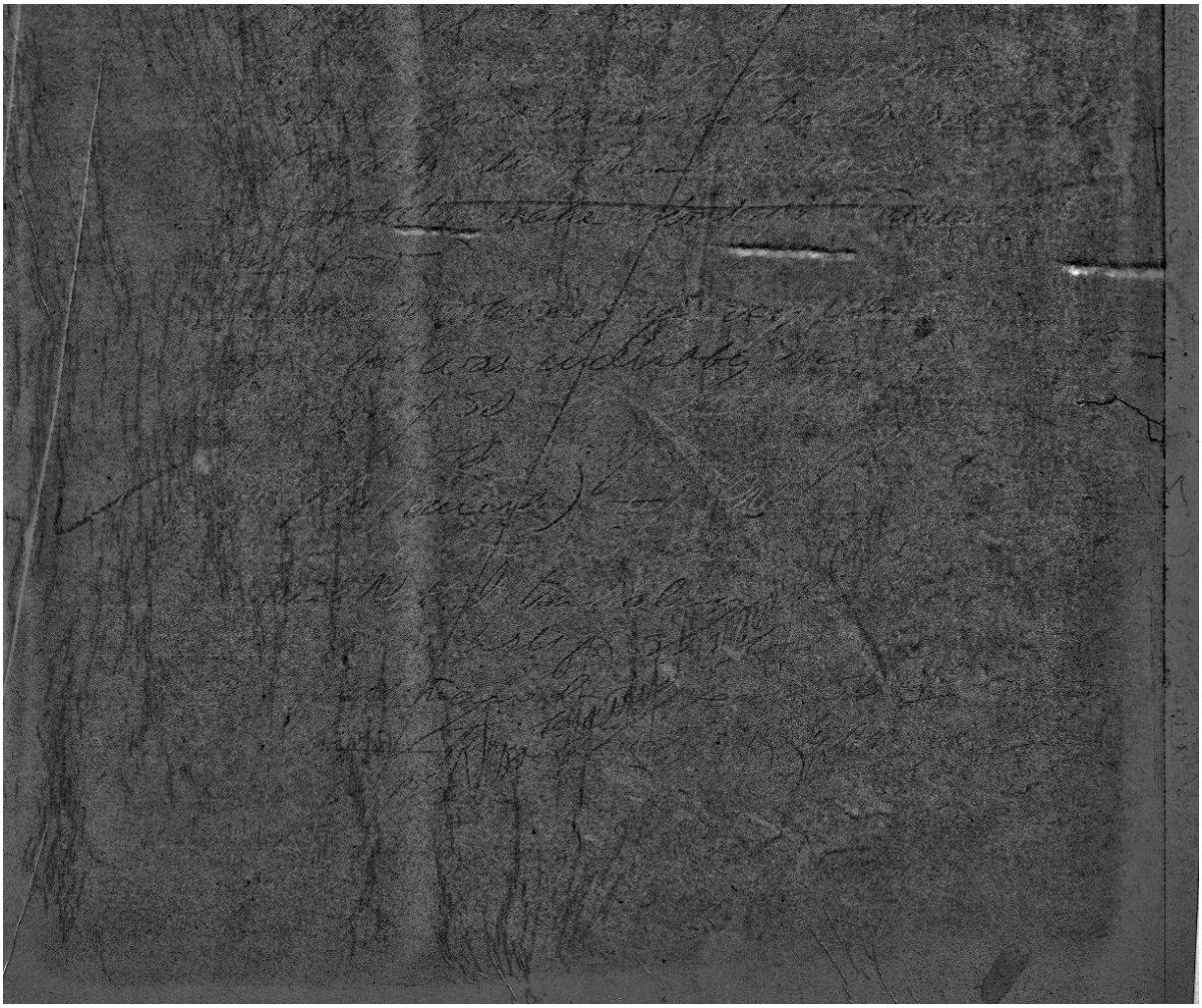


Figure 7. Close-up of Page 2 Lower Region with Indented Writing – Layer Subtraction

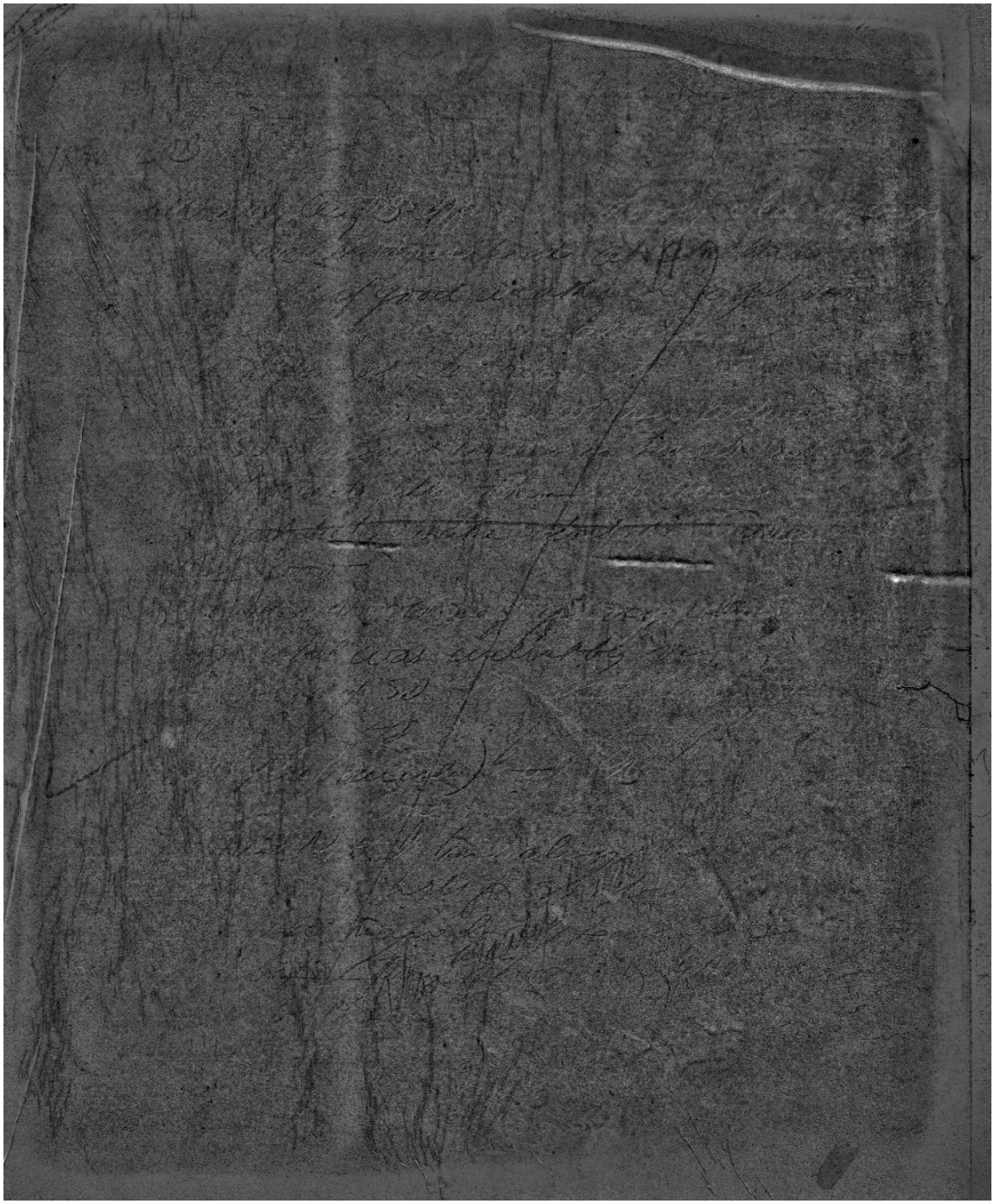


Figure 8. Close-up of Page 2 Lower Region with Indented Writing: Trans, Darkened Overlay

Figure 9. Close-up of Page 2 Lower Region with Indented Writing – Contrast Adjustment

If further analytical work was to be undertaken to assess options for more effective rendering of potential indented writing, the following option of optical profiles might be a possibility. White-light optical profilers are used to measure surface information in a wide variety of applications, including for optical glass, ceramics, microelectromechanical systems (MEMS), etc.

The WYKO NT1100 benchtop optical profiler system is capable of measuring surface profiles with $\Delta z \leq 1$ mm over areas of the order of a few square millimeters, and it is possible to stitch several such renderings together in software (before display) to generate mosaics of larger areas. The best measurements on this system are obtained from specular features. This means that the sensitivity of the measurement might be improved if a wetting agent is applied to the paper, or humidification of the paper to increase the water content.

An experiment on a test sample of indented writing (not the document under examination) showed the possibility and potential for utilizing an optical profiling technique to render indented writing as an addition to the current ESDA technique. White-light interferometric optical profilers may be a useful tool for mapping the surface topography of paper in forensic applications**.**

Spectral Curves

Comparison of Ink on Imaged Pages

The following spectral curves illustrate the unique spectral response based on the chemical composition of the inks on each page. A statistically significant range of points from each of the inks present on every page imaged under the hyperspectral imaging system was captured and the below curves represent the average of these points for each ink. Curves were generated in the ENVI-ITT software package and demonstrate the *x,y,z* coordinates for the "pixel" sample of the ink taken. The location on the Cartesian coordinate for the page are represented by *x* and *y*, while *z* represents the specific spectral response for that ink throughout all wavelengths, from 365-1050nm.

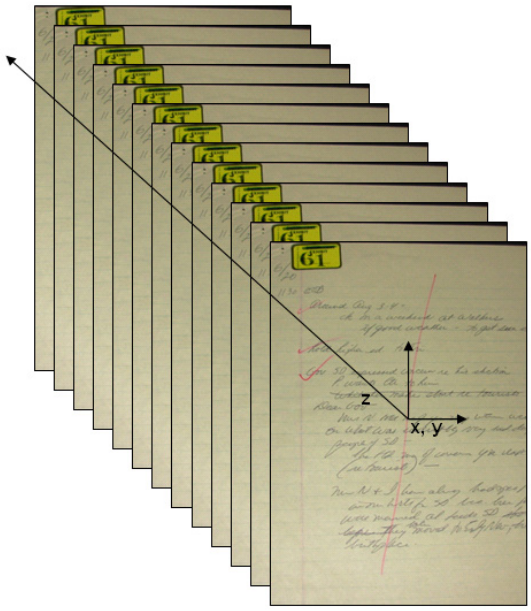


Figure 10. *X, Y, Z* Positioning for Pixel Sample of Ink

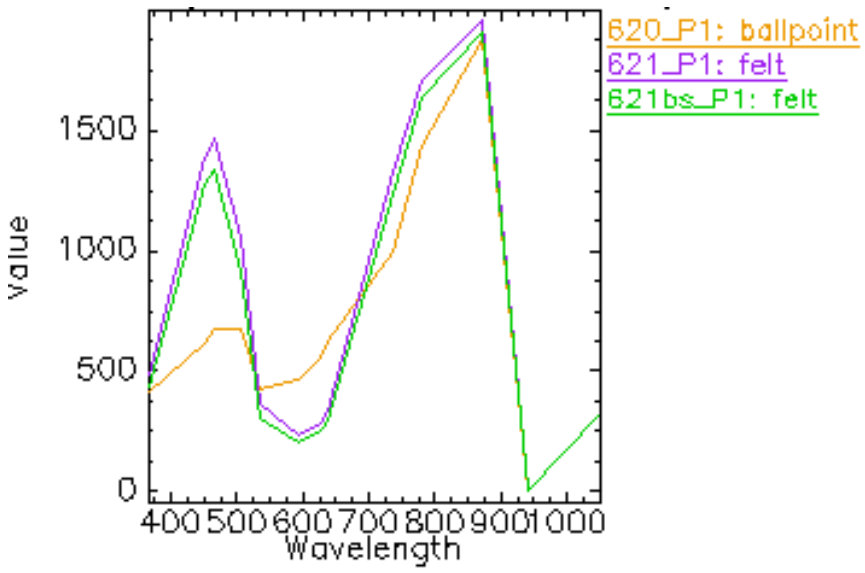


Figure 11. Comparison of 6/20, 6/21 Text Inks

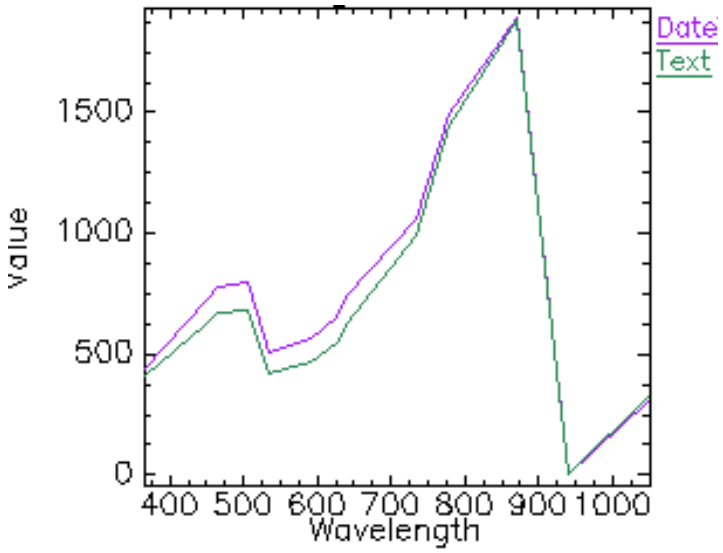


Figure 12. 6/20 Page 1 - Comparison of Inks: Date, Main Text

Small differences between the body of the text on the page and the date (page 1) and page number (page 2) are illustrated in these two figures.

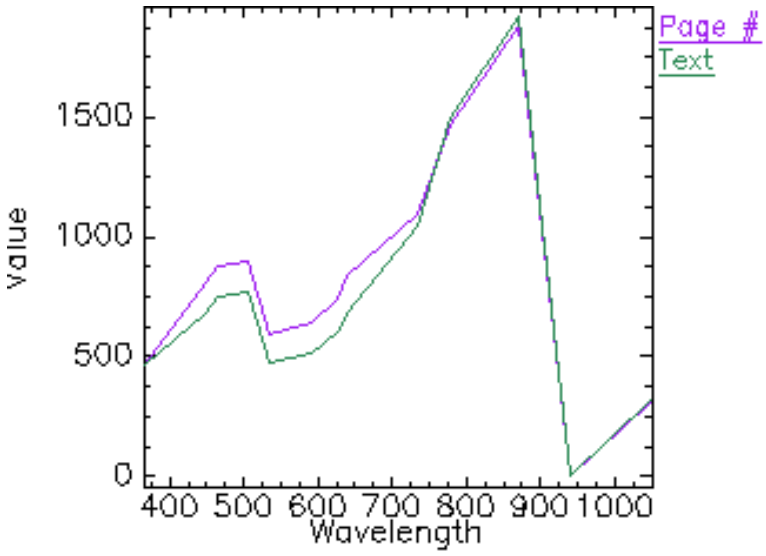


Figure 13. 6/20 Page 2 - Comparison of Inks: Page No., Main Text

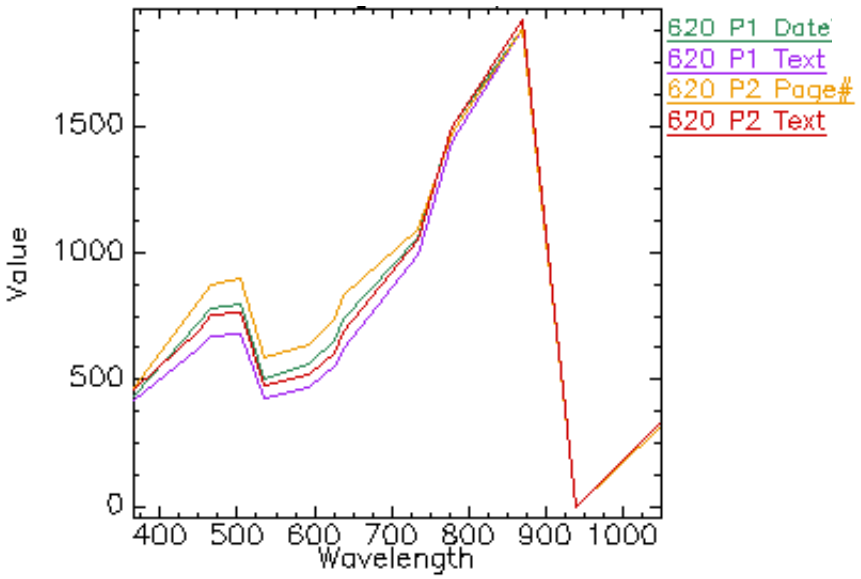


Figure 14. 6/20 Page 1 and 2 - Comparison of Inks: Date, Page #, Main Text

No apparent difference for the date and body of the text for 6/21 page 1 was perceived with this technique.

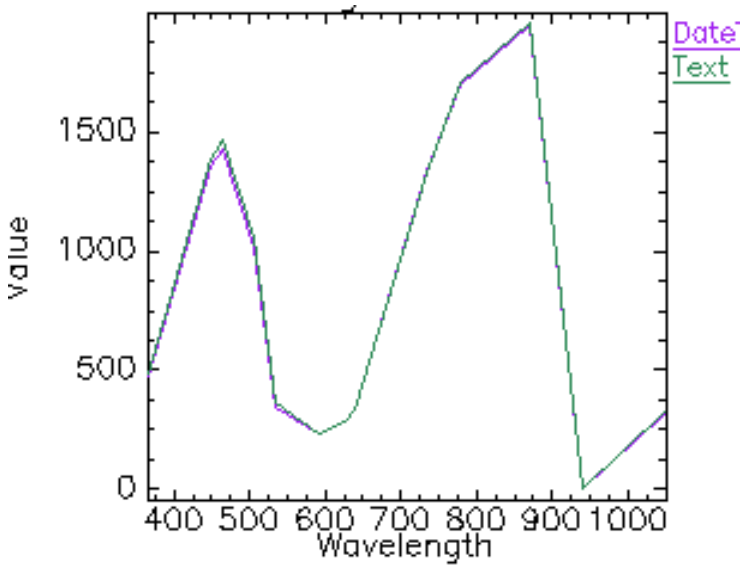


Figure 15. 6/21 1330 Page 1 - Comparison of Inks: Date, Main Text

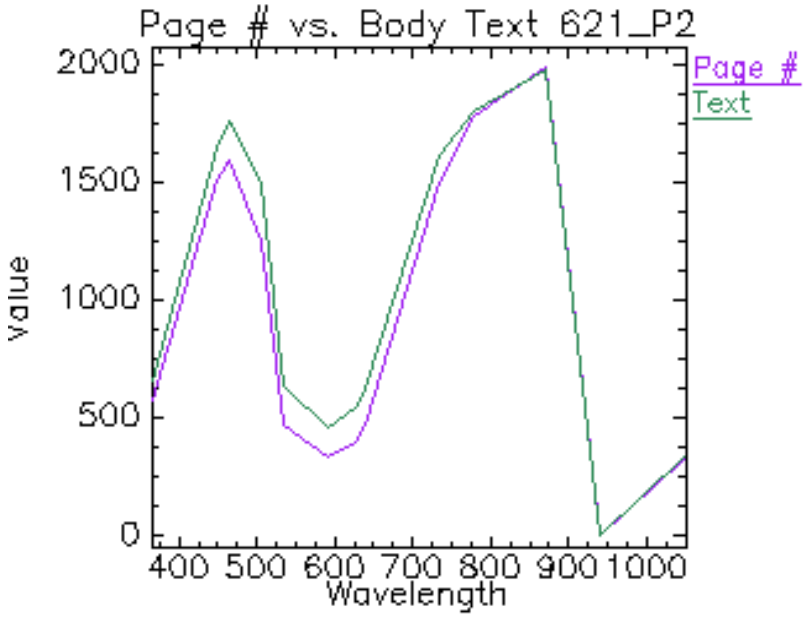


Figure 16. 6/21 1330 Page 2 - Comparison of Inks: Page #, Main Text

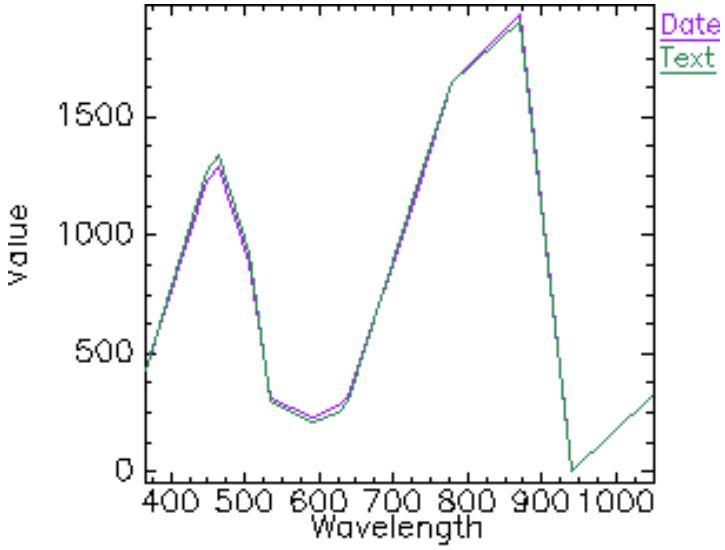


Figure 17. 6/21 Be sure - Comparison of Inks: Date, Main Text

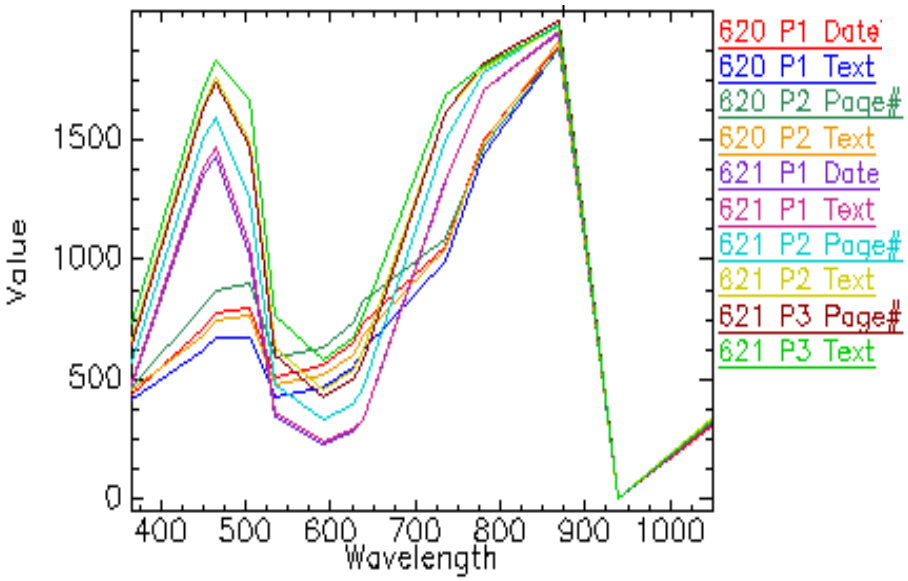


Figure 18. Comparison of 6/20, 6/21 All Inks

The below figure illustrate a comparison of all pages for the red "strike-through" ink used on all pages. Some difference observed may be due to varying intensity of the strike-through. Samples of both "ticks" and strike-through lines were assessed.

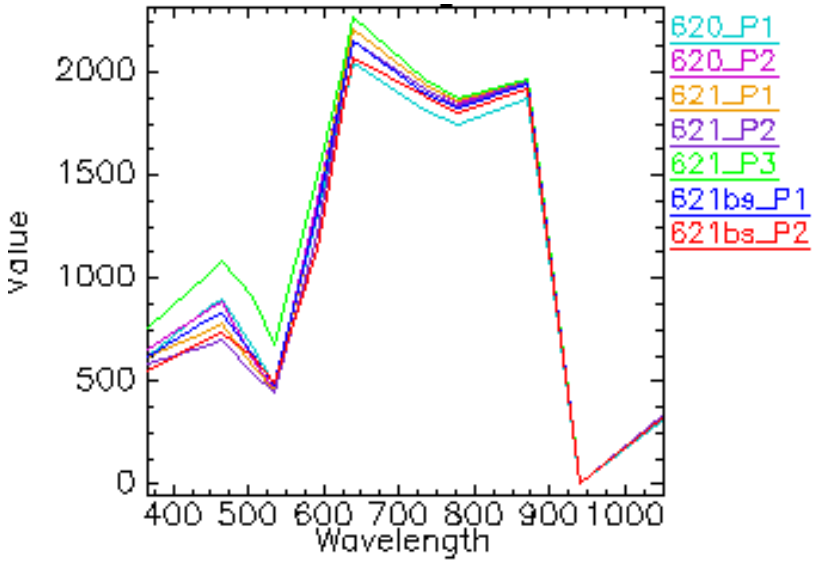


Figure 19. Comparison of 6/20, 6/21 Red Inks

Comparison of Paper Substrates of Imaged Materials

Comparisons of all pages showed remarkable similarity for all pages suggesting that the paper is either from the same manufactured series of notepads within a specific timeframe, or that the manufacturer was capable of maintaining a consistent product over an extended period of the production of these notepads.

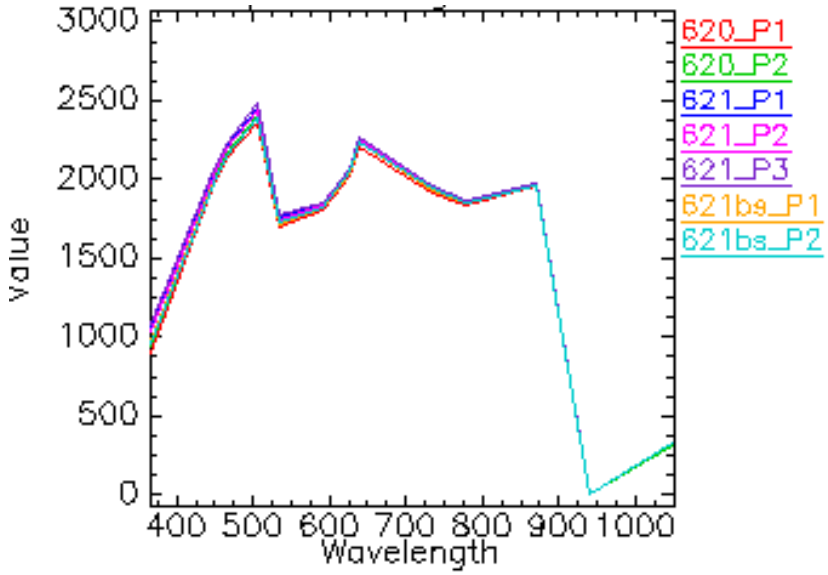


Figure 20. Comparison of 6/20, 6/21 Paper Substrates

Advanced Processing for Additional Information

The technique of hyperspectral imaging allows the utilization and capture of the unique spectral response of all materials present in the document being images (both substrate – paper, and media – inks) to capture any additional information. Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation that converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component accounts for as much of the variability in the data as possible and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (uncorrelated to) the preceding components.

The following images illustrate principle component analysis (PCA) images rendered in pseudocolor to enhance differences and reveal previously undetected features. Pseudocolor combinations most likely to render differences were chosen.

Fingerprints

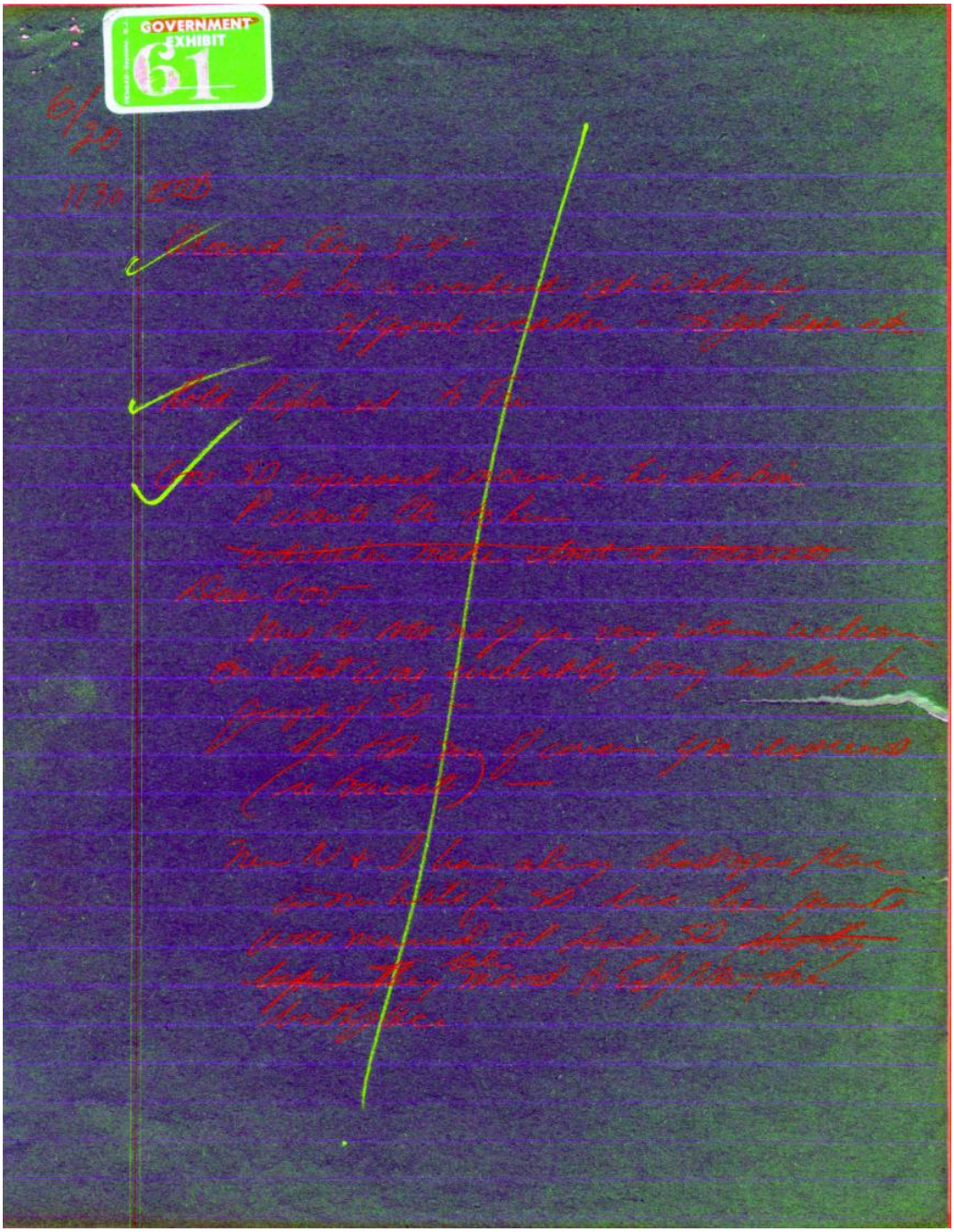


Figure 21. 6/20 Page 1 (recto) – PCA Pseudocolor Image of Fingerprint in Center of LHS Margin

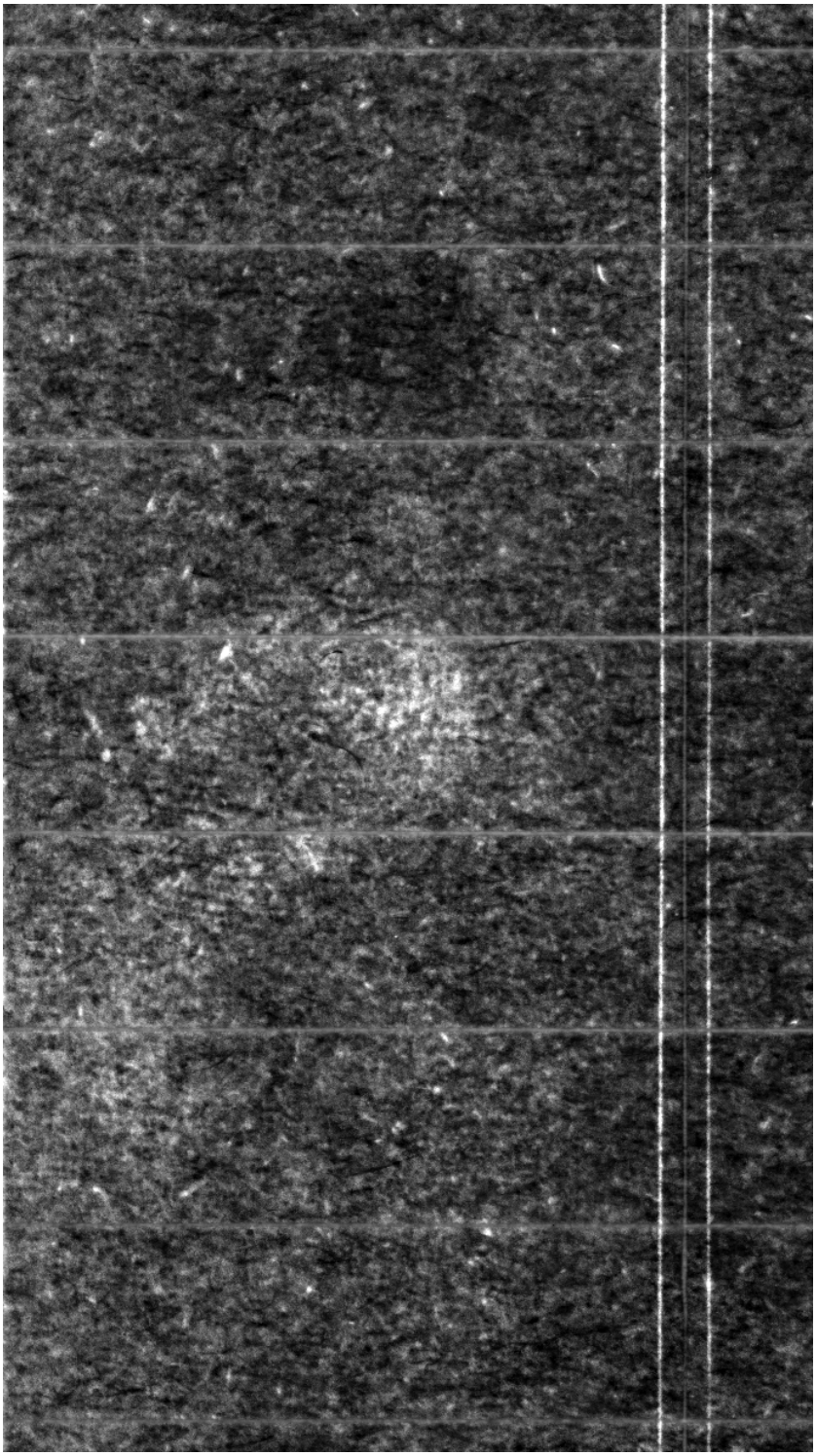


Figure 22. 6/20 Page 2 (recto) – Fingerprint: Magnified Grey-scale Image

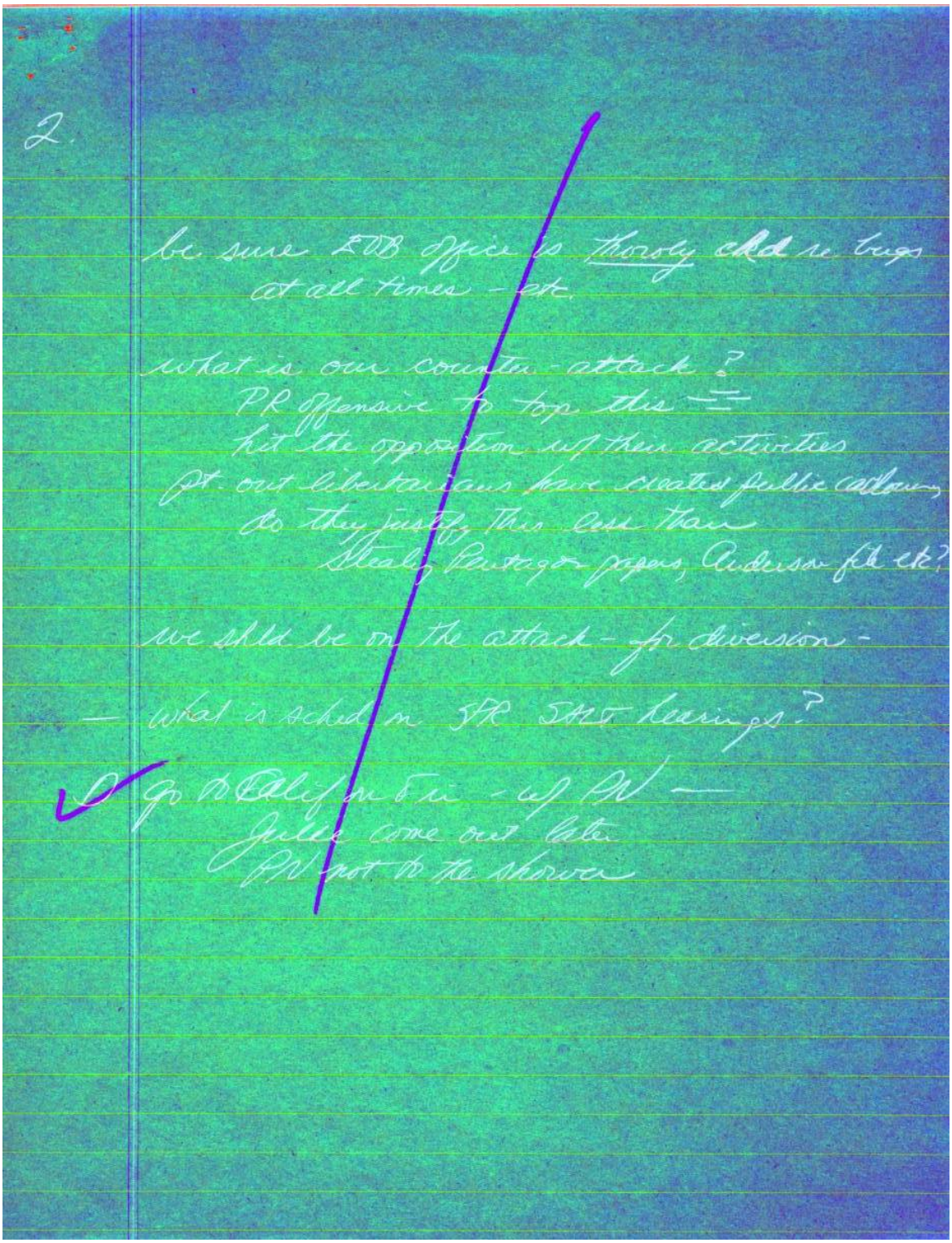


Figure 23. 6/20 Page 2 (recto) – PCA Pseudocolor Image of Fingerprint in Center of LHS Margin

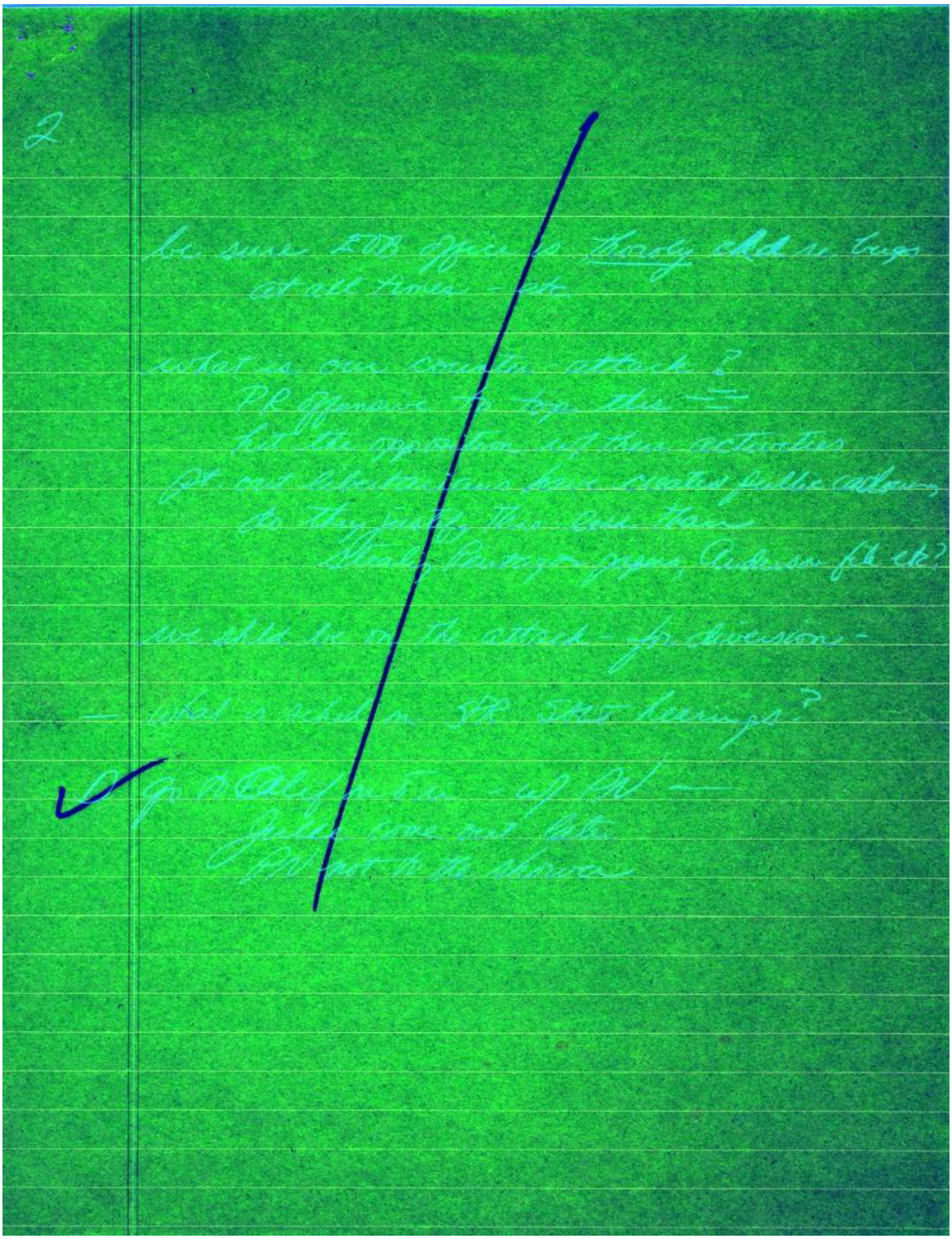


Figure 24. 6/20 Page 2 (recto) – PCA Pseudocolor Image of Fingerprint in Center of LHS Margin

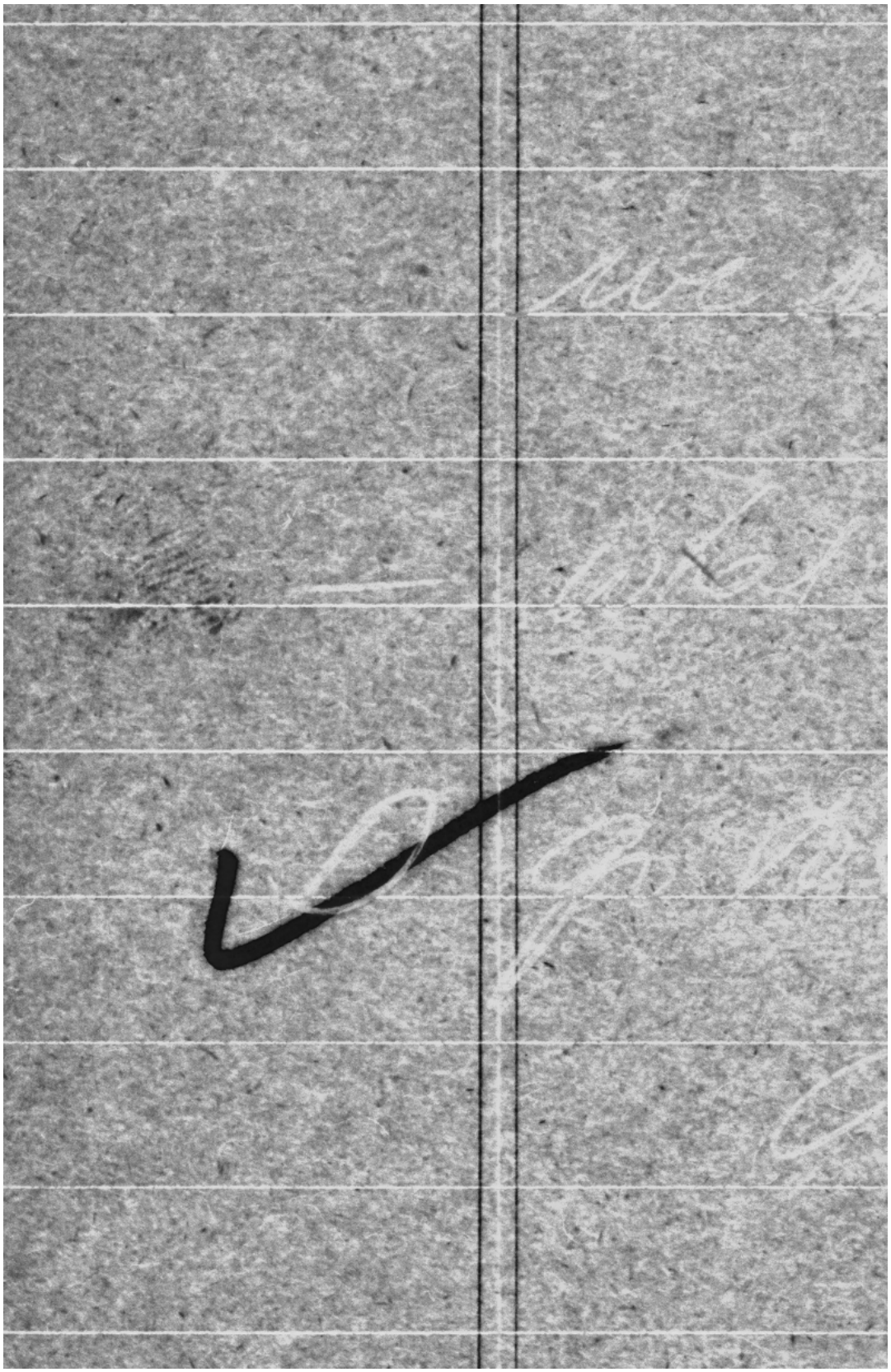


Figure 25. 6/20 Page 2 (verso) – Fingerprint: Magnified Grey-scale Image

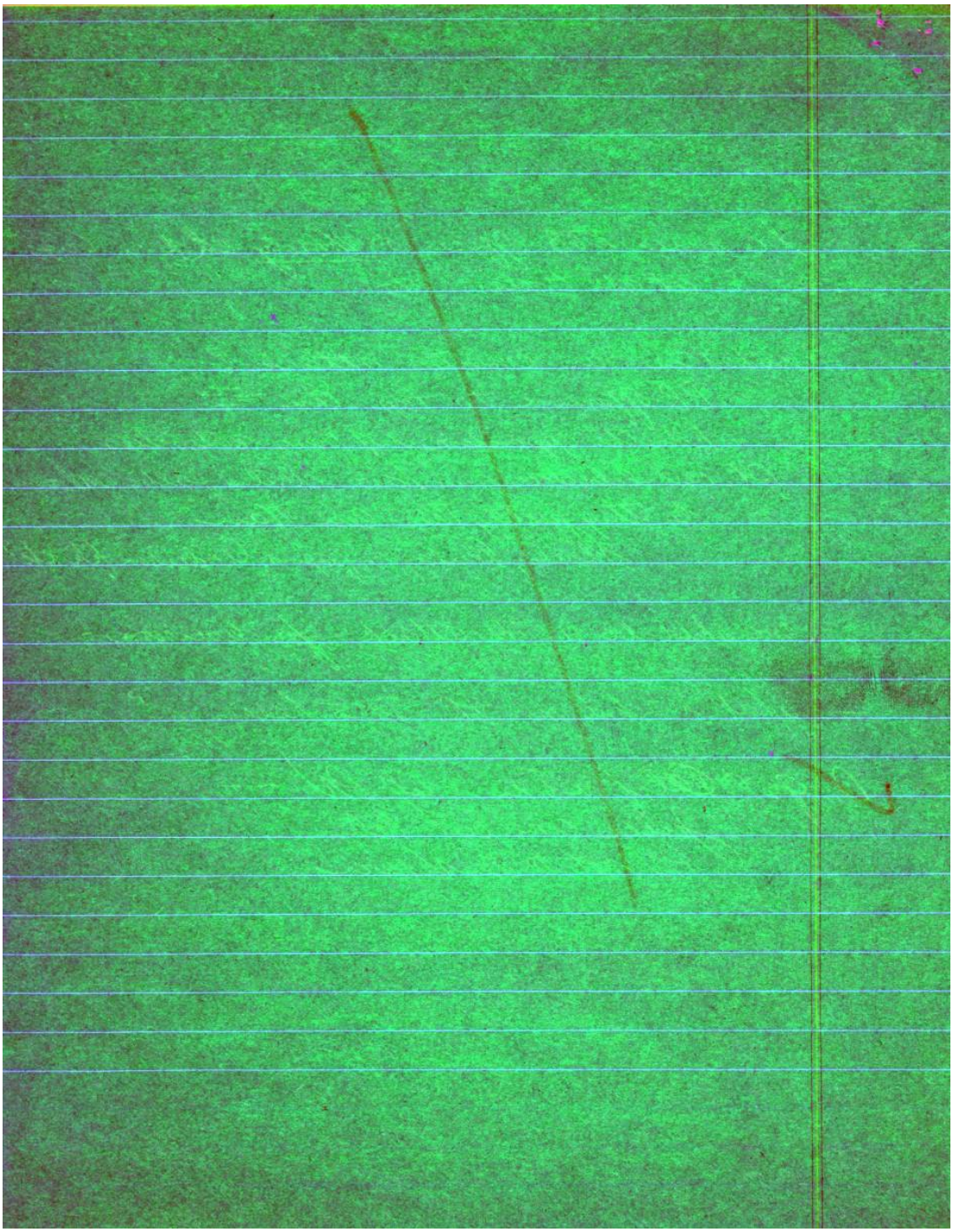


Figure 26. 6/20 Page 2 (verso) – PCA Pseudocolor Image of Fingerprints in Center of RHS Margin

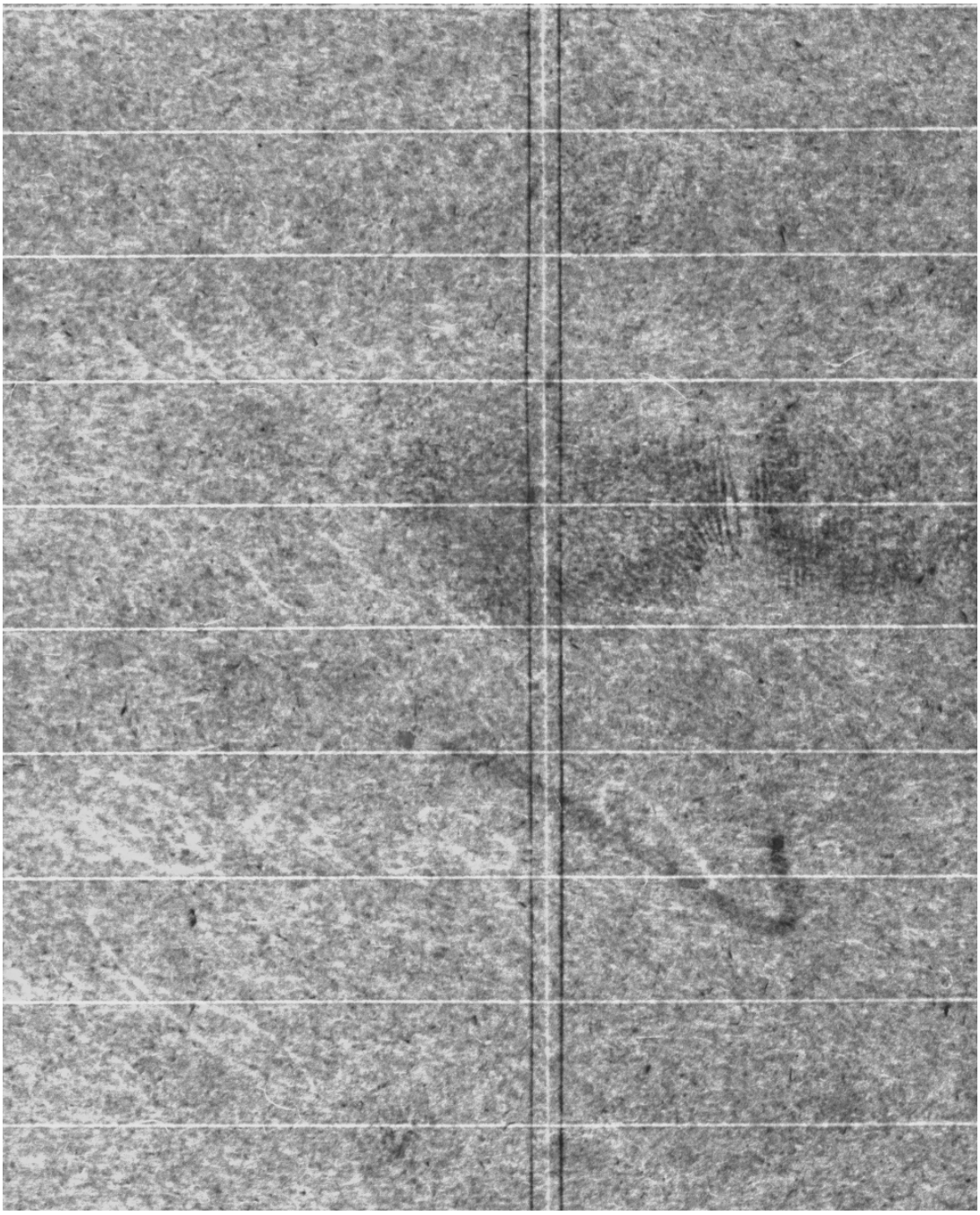


Figure 27. 6/20 Page 2 (verso) – Fingerprint: Magnified Grey-scale Image

Handwriting Emphasis

Hyperspectral imaging can be utilized to assess handwriting. An example of this is shown below, where PCA and pseudocolor image processing can reveal characteristics of writing pressure and emphasis.

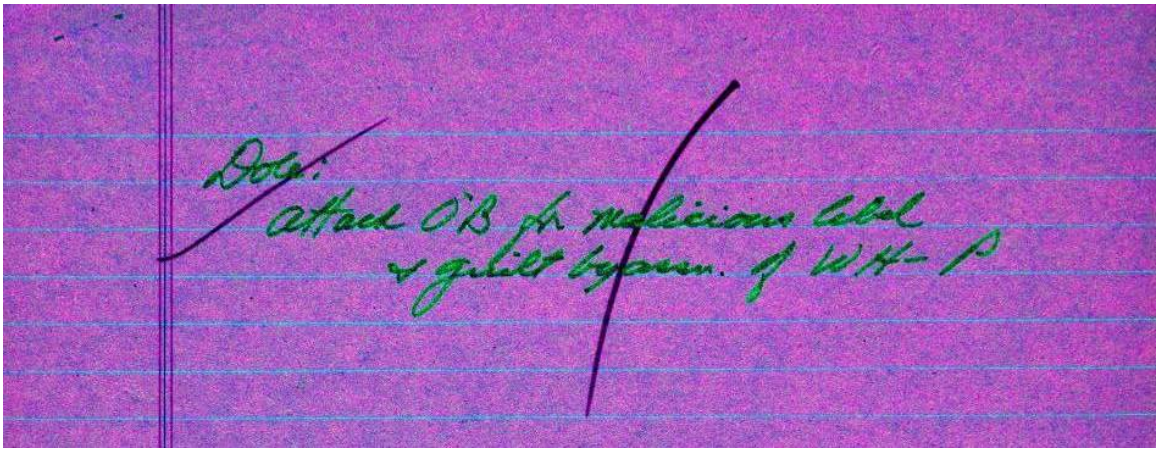


Figure 28. 6/21 Be sure 10 band PCA [2-11] Pseudocolor Image

Alignment of Staples

An examination was undertaken from overlaying all pages of 6/20 to determine whether it was possible to determine what orientation the staples were in in relation to the tears on the RHS of these two pages. As shown below, page 1 and 2 of 6/20 were overlaid.

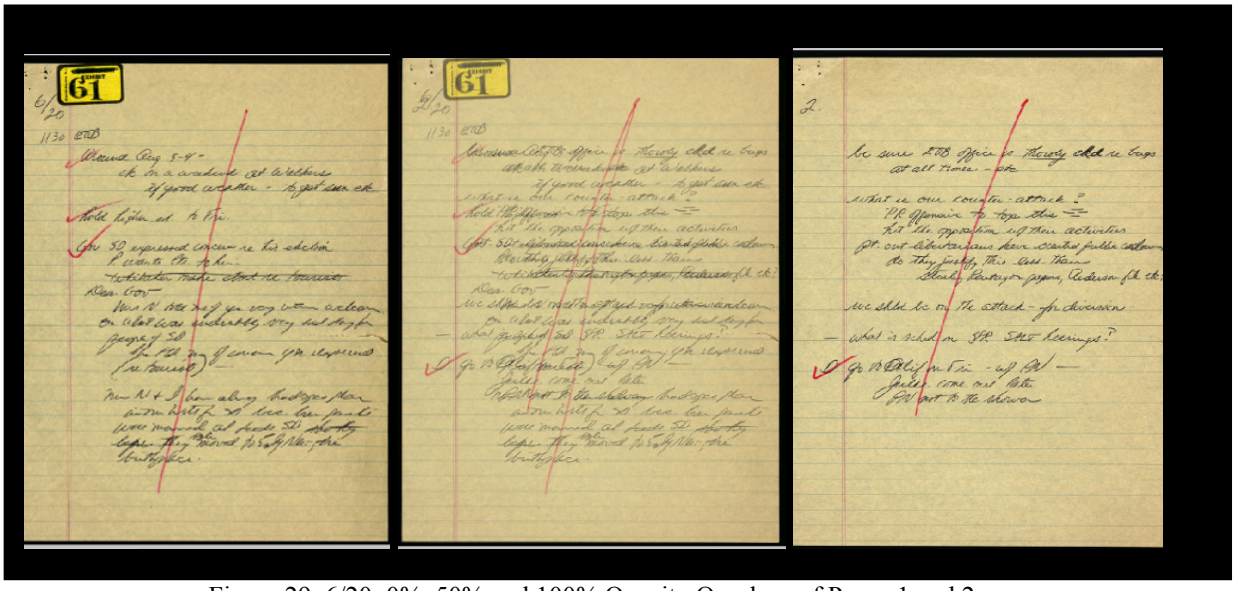


Figure 29. 6/20: 0%, 50% and 100% Opacity Overlays of Pages 1 and 2

Aligning page 2 to page 1 of 6/20 to assess whether the tear was created when one set of staples were in place, illustrated that the alignment to match the bottommost pair of holes in both pages matched best for the holes on the 45-degree diagonal. These appear to give the best match of the right-hand margin tears.

Additional Processing
Evaluation was undertaken of the PCA of all four sides of pages 6/20, pages 1 and 2 (recto and verso) using statistics from the yellow paper (with lines), and the yellow paper only (no page lines, handwriting, red lines, check marks, labels etc.) The statistics were calculated and then applied to the entire image to make 13 PCA bands.

Examination of 6/20, page 1 revealed the following: band 12 of this Principal Component Analysis outlined two red lines. The second red line was a "ghost" type image of the strike-through red line from page 2 illustrating evidence of some transfer from page 2 to page 1. When the same technique was applied to page 2, no additional red felt lines were apparent; showing no transfer from an adjacent page. This lack of transfer could suggest that there were no additional pages that were transcribed on the same day.

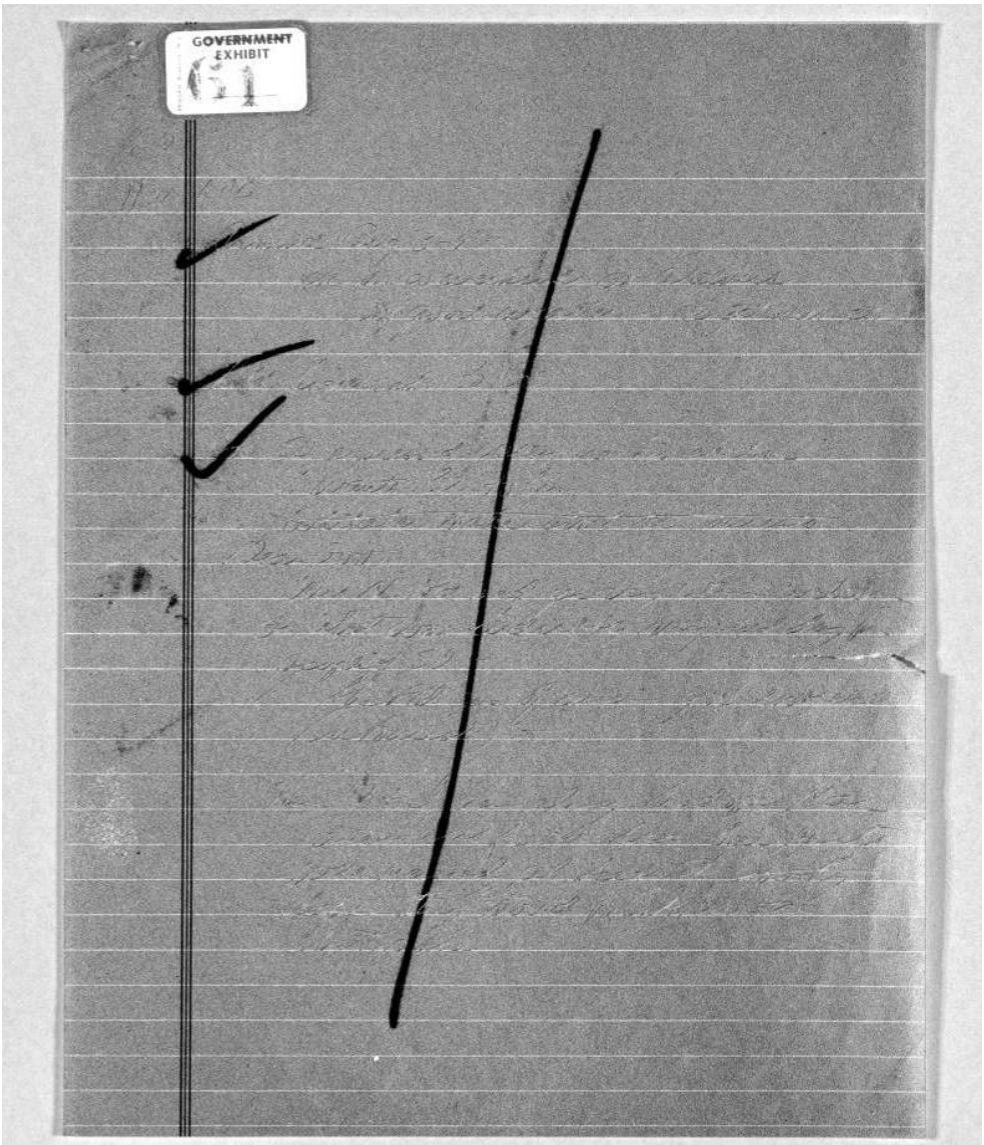


Figure 30. 6/20 1130 Page 1 recto: PCA Image from Masked Statistics (paper with no lines)

Figure 31 illustrates another fingerprint not previously visible through other imaging. Four lines above the red check mark is the previous fingerprint rendered more obvious in black ink; while immediately below the faded checkmark is a newly rendered "white" fingerprint visible with this processing technique.

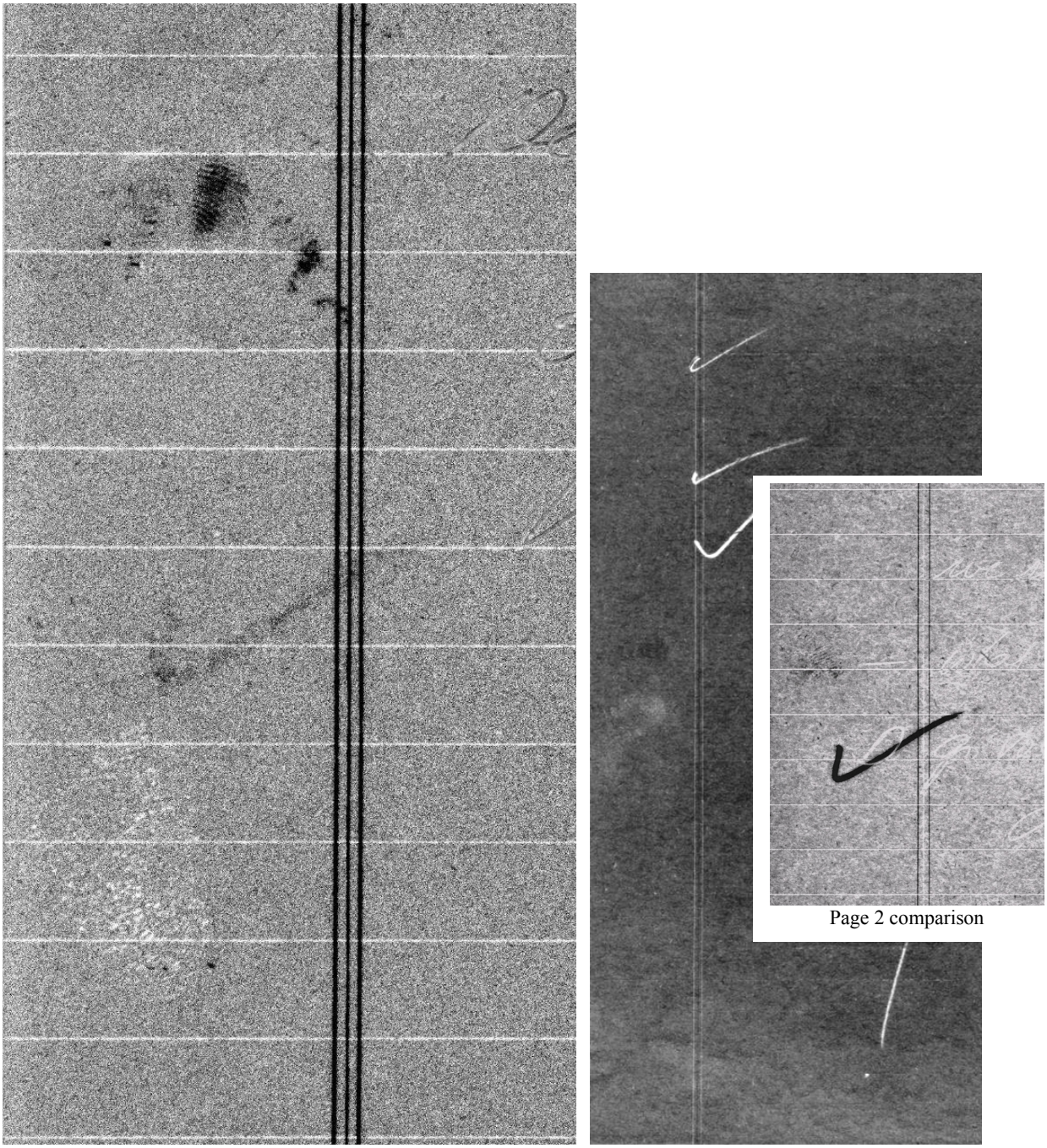


Figure 31. 6/20 1130 Page 1: PCA Image from Masked Statistics (paper with no lines) Magnified Image of Additional Fingerprint from Page 2 recto *(inserts of previously illustrated prints)*

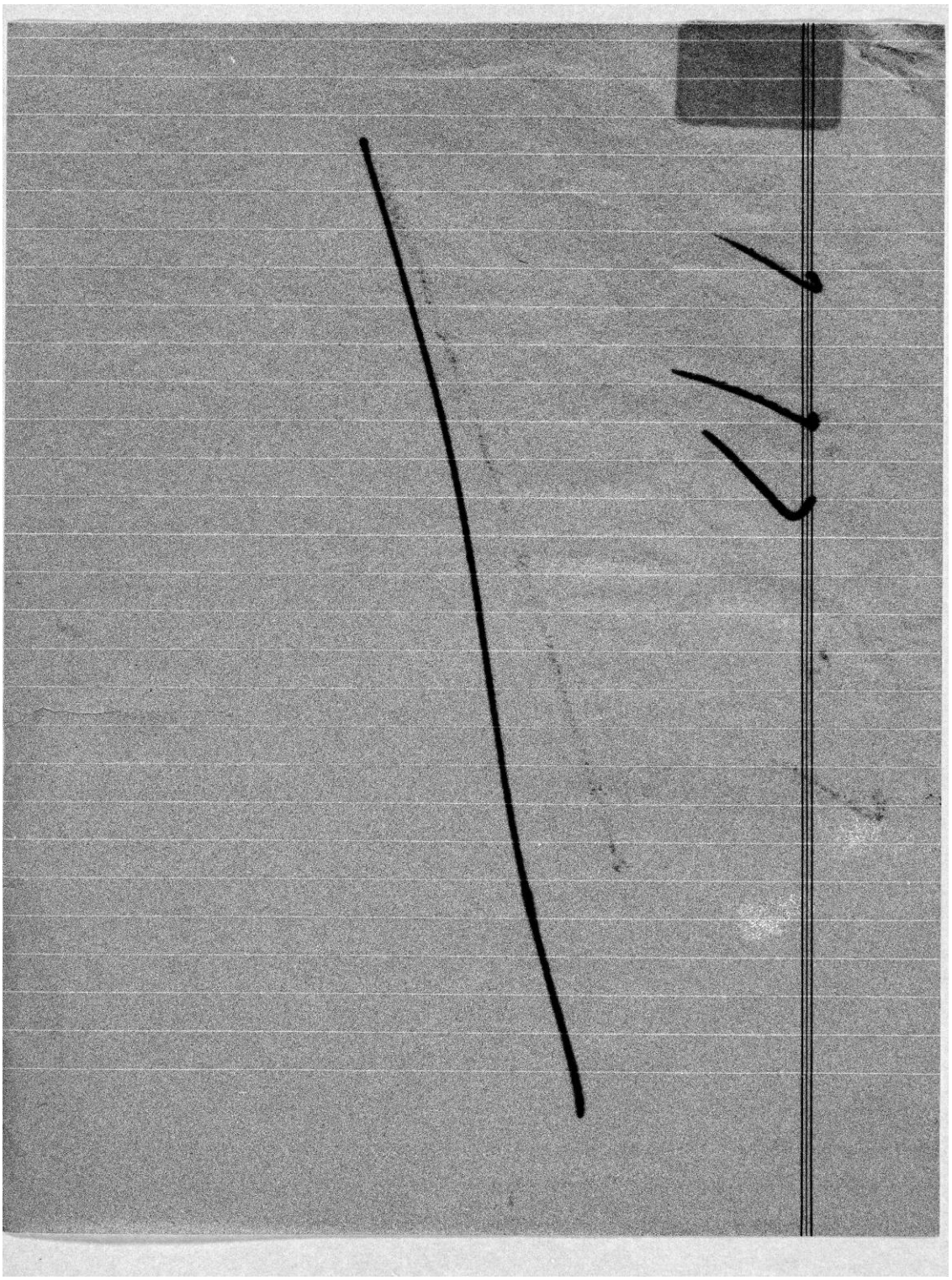
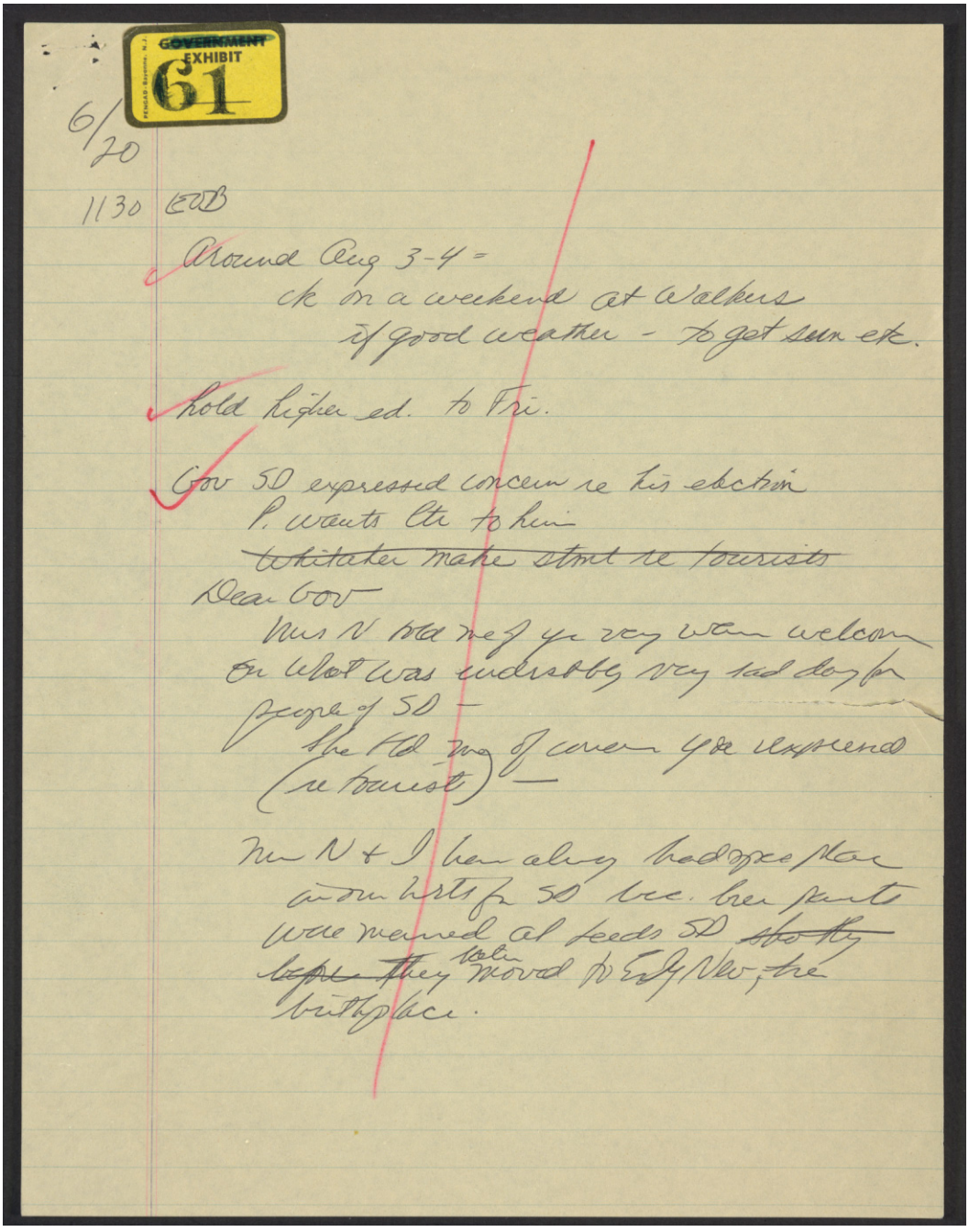
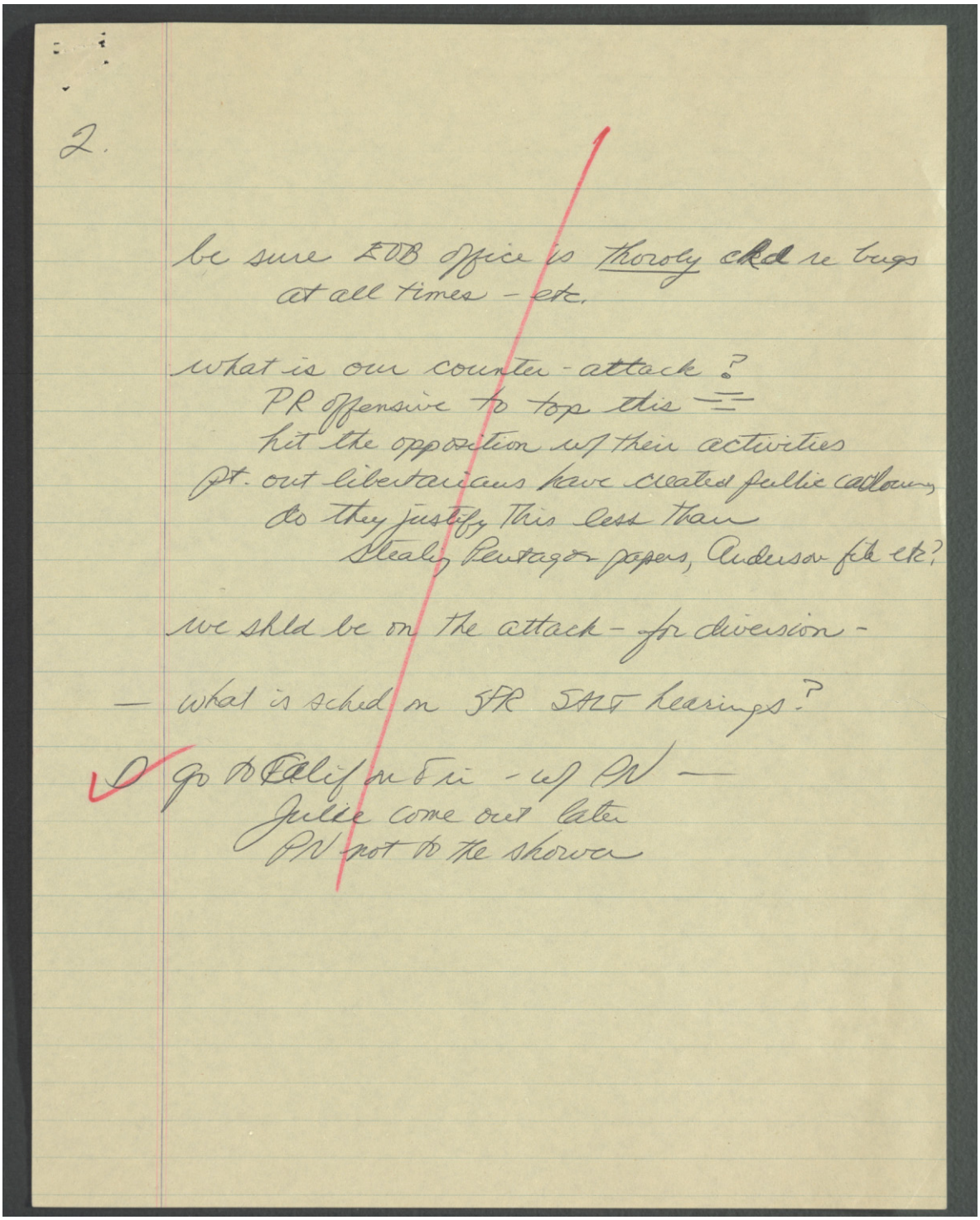


Figure 32. 6/20 1130 Page 1 verso: PCA Image from Masked Statistics (paper with no lines)

Full Spectrum Color Images

Full spectrum color images were created for all pages imaged by the hyperspectral imaging system. Using a special algorithm and the precise spectrometric color response from the Gretag Macbeth color Checker, an L*a*b* color space is created from the seven wavebands and monochrome images within the visible region. Since normal cameras have only three color channels – red, green, blue (RGB) – this imaging process can create a high resolution image that closely resembles the original document. Examples of these high resolution images are illustrated below.





2.

be sure EOB office is thoroly chkd re bugs
at all times - etc.

what is our counter-attack?
PR offensive to top this ==
hit the opposition w/ their activities
pt. out libertarians have created public outcry
do they justify this less than
Ellsberg Pentagon papers, Anderson file etc?

we shld be on the attack - for diversion -

— what is sched on SR SALT hearings?

✓ go to California - w/ PN —
Julie come out later
PN pot to the shower

Conclusions

Research in the above sections outline the usefulness of the application of hyperspectral imaging for non-invasive analysis to answer a range of questions and issues in relation to the examination of the Haldeman papers from 6/20 and 6/21/1972.

Overlays of indented ESDA and HSI images provided some success at increasing the contrast to enhance readability of indented writing observed after the ESDA examination. White-light interferometric optical profilers could potentially be applied to hyperspectral images to enhance the readability of indented writing and other hidden features.

Spectral curves obtained from the unique chemical response of compounds present in the spectral images indicated some differences between the text on the body of the page and the top corner margin page number and date text for the two pages from 6/20. No observable difference was apparent for the paper substrate used for the pages examined from 6/20 and 6/21.

Additional processing using techniques such as PCA and pseudocolor processing enhanced and revealed the presence of finger prints on some of the pages examined. A sample of how the intensity of pressure in handwriting can be examined was included.

Further processing that included masking regions to take the statistics from the background paper (with and without lines and other features) revealed previously unseen features including the transfer of the red “strike-through” line from page 2 to page 1 of 6/20, as well as another fingerprint not detected with other techniques.

Full spectrum color images were created for all pages imaged by the hyperspectral imaging system to create high resolution images that closely resemble the color and features of the original document.

References

C. Fisher, and I. Kakoulli, *Reviews in Conservation,* **7**, 3-16 (2006).

F.G. France, D. Emery, and M.B. Toth, *Library Quarterly: Digital Convergence: Libraries, Archives and Museums in the Information Age*, **80(1)**, 33-59 (2010).

W.A. Christens-Barry, K. Boydston, F.G. France, K.T. Knox, R.L. Easton, and M.B. Toth, *Proc. SPIE, San Jose, CA*, **7249(8)**, 1-10 (2009).

W.A. Christens-Barry, K. Boydston, and R.L. Easton., Jr, *Proc. SPIE, San Jose, CA*, **7528**, 0M (2010).

F.G. France, W.A. Christens-Barry, M.B. Toth, and K. Boydston, *Proc. SPIE, San Jose, CA*, **7531-13**, 1-12 (2010).

R.L. Easton., Jr, K.T. Knox, W.A. Christens-Barry, K. Boydston, M.B. Toth, D. Emery, and W. Noel, *Proc. SPIE, San Jose, CA*, **7531**, 0D1-11 (2010).

F.G. France, *Eikonopoiia: Symposium on Digital Imaging of Ancient Textual Heritage, Helsinki, Finland*, 51-64 (2010).

D. Emery, F.G. France, and M.B. Toth, *Proc. Of IS&T: Archiving 2009*, 137-141 (2009).