https://doi.org/10.2352/issn.2168-3204.2018.1.0.32

Multispectral Imaging for Scientific Analysis and Preservation of Cultural Heritage Materials

Meghan Wilson;Fenella France, and Chris Bolser; Library of Congress; Washington, District of Columbia, USA

Abstract

Multispectral imaging is a digital imaging technique that adds depth to understanding cultural heritage collections. When adhering to standards and best practices it can afford a scientific analysis with commensurate integrity. The Library of Congress was one of the first institutions to implement this technology in their primary workflow as a standard for examination and preservation of its collection items. The Preservation Research and Testing Division (PRTD) has spent the past decade focusing on the development of standards and procedures for this imaging technology while also expanding its applications. Additionally, the Library of Congress has taken initiative in adapting their rigorous methodologies for practical integration of spectral imaging at other institutions. This technique expands the traditional concept of an image, while retaining the precision required for accuracy of reproducibility.

Introduction

The Preservation Research and Testing Division is a unique department to be found within a library or archive. Its capacity for dedicated scientific research with access to an expansive and diverse collection of cultural heritage materials has enabled the ability to focus on the development of multispectral imaging (MSI) workflows, standards, and practices for daily use and the expansion of imaging technology applications for scientific analysis.

The added value of including spectral data in an image has redefined documentation for preservation, but requires a sustained focus on standardization of process and procedures. The increased demand for timely and non-invasive preservation investigations created impetus for expanding what could be captured to answer preservation questions, while enhancing knowledge of collection materials.

Defining Multispectral Imaging and System Specifications

Multispectral imaging is a non-invasive technique that captures information at specific wavelengths along the electromagnetic spectrum, including those beyond the visible range. This allows for spectral characterization of inks, colorants, substrates, treatments, assessment of environmental factors, and recovery of obscured or deteriorated content that is not visible in natural light.

The Library of Congress's standardized spectral imaging system is comprised of a camera and illumination panels with integrated software and hardware. The 50 megapixel monochrome camera uses a CCD digital back with 8176 x 6132 pixel array at 6.0 micron pixel size and a 120mm Copal-0 hyperspectral lens. Attached is a dual filter wheel containing (ultraviolet (UV) pass, UV block, Red, Orange, Green, and Blue) optical filters to increase the range of captured information to include fluorescence emission and UV reflectance images. Illumination in reflected, low angle raking (or side lighting at a 15^{o} angle), and transmitted modes is provided by integrated panels containing light emitting diodes (LEDs) spanning 365 nanometers (nm) to 910nm. LEDs produce low amounts of light and heat ideal for working with sensitive cultural heritage materials. Integrated image capture software controls the digital camera back, filters, and illumination system, including the ability to adjust the exposure duration of each individual LED to enable optimal histograms in each image of the sequence. All captured images are fully registered, alleviating the need to adjust each image cube for pixel shift and allowing for a streamlined transition between image capture and processing.

The need to recreate and reproduce conditions for documentation and tracking change over time drove the careful development and standardization of the imaging process, assuring that all imaging modalities were controlled and did not contribute to any observed change in the collection materials being imaged.

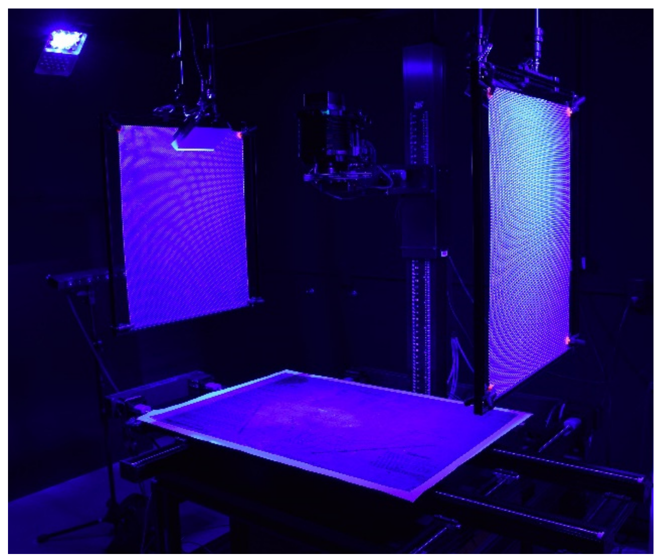

Figure 1. Multispectral imaging system at the Library of Congress

Multispectral Imaging Standards

For the past ten years the Library of Congress has been evolving multispectral imaging technology for application of preservation research analysis in a cultural heritage institution. This particular spectral system was initially developed by adapting components from a range of imaging fields including medical and geospatial. Therefore its workflow requires versatility, pulling from established standards and pioneering their integration to create a unified process [1].

Image quality is the first practical standardization criteria. This of course includes static parameters like focus (a pixel is either in focus or it's not) but also more flexible components like resolution, a choice which is generally at the discretion of the photographer. The parameters for variables like this largely conform to established digital photography standards [2]. While PRTD's system has the capacity to image at resolutions between 400 pixels per inch (ppi) and 4000ppi, all objects are baseline imaged at a standard of 600ppi with additional resolutions used to meet the goals of the project as needed. To maintain this quality of resolution, objects that are larger in size may extend past the field of view of the camera, requiring the need to image their entirety in multiple tiles. After examining the capabilities of the Copal-0 hyperspectral lens, PRTD chose to image at a standard aperture of F11 to maintain sharp focus to the edge of the image, especially critical in those instances when the object takes up the totality of the camera frame.

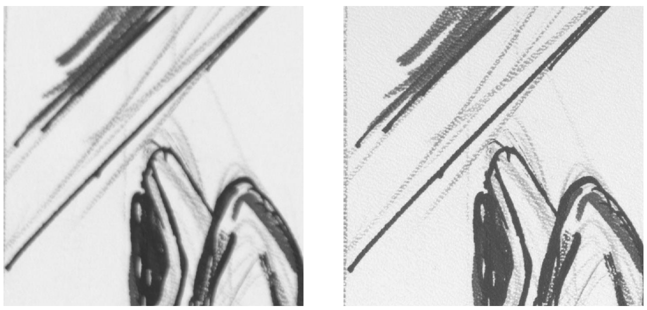

Figure 2. Comparison of sharpness at the edge of an image taken at F8 (left) and F11 (right)

Image processing of multispectral data requires the inclusion of color targets and white reference standards in the imaging scene during capture. PRTD dedicates a color checker for use only with the multispectral system, otherwise keeping it safely covered and stored to limit exposure to other harsher light sources like continuous fluorescent room light. This prevents fading and subsequent negative impact on image processing, especially change over time analyses. PRTD uses a separate square of Spectralon, a fluoropolymer with the highest diffuse reflectance in the UV-VIS-IR spectrum, as the main white reference standard which is stored similarly to the color checker.

Flat field correction is a critical step in a multispectral imaging workflow. Flat fields compensate for lack of lighting uniformity. They should be captured from a smooth material of uniform reflectance and under identical camera and illumination conditions and settings with a new set of flats captured for any changes in those parameters.

Standardizing imaging variables requires routine measurement and calibration of the imaging equipment to ensure the integrity and repeatability of imaging data for scientific use. PRTD runs an annual equipment master check of its multispectral imaging system. This includes most importantly confirming the power output, peak wavelength, and full width at half maximum of each LED as well as measuring each patch of the color target and Spectralon to verify no shift has occurred in their spectral response.

Consideration for digital preservation of data has been key in the creation of the Library's current standards. Specific metadata requirements are critical for the longevity of data. The nature of multispectral imaging requires a conglomeration of particular types of metadata that no single standard encompasses. The system must retain technical information pertaining to the capture and unique illumination parameters, while allowing for standard catalog and bibliographical information about the object being imaged, and it is likewise imperative to incorporate identification of image elements like spatial information linking the stack of registered captured images to one another. The software used for multispectral capture at the Library of Congress includes a metadata table with entry fields customized for cultural heritage imaging. These fields are based on Dublin Core metadata terms [3]. They allow for bibliographic entries to link to the object's original record, information about the material composition of the object, and parameters about the imaging set up to enable repeatability which is especially important for long term research projects that examine materials over time. This descriptive metadata is then embedded into each image capture along with the technical data and is readable through exchangeable image file (exif) tags.

The use of sustainable file formats is additionally imperative for long-term digital integrity. The Library's multispectral data is created in digital negative graphics (.DNG) and 16-bit tagged image file format (.TIFF) formats. The lossless TIFF format is crucial for image processing as it can be edited and resaved with no loss in image quality. For this reason, the .JPG format is avoided as it compresses image files, resulting in a loss of data. While the reduced size of .JPG files can be convenient for quickly sharing images, they are only created as derivative images from the original archive quality .TIFF. Proprietary file formats are also avoided as these cause significant restrictions on the number of software programs that can read and open them. This is cumbersome when sharing data and is especially problematic for archiving purposes as future access of a file in a format readable only by a software no longer in existence could render the file defunct.

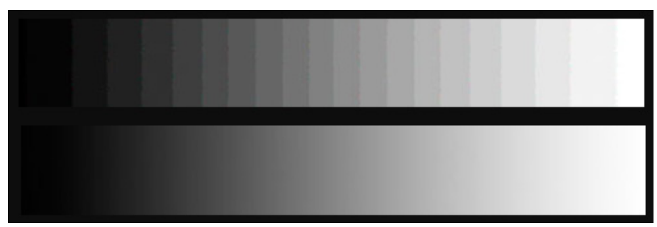

Figure 3. The impact of compression and poor bit-depth (top)

A strong interoperable data management infrastructure is crucial for both the archiving and accessibility of comprehensive multispectral image cubes and derived scientific data. This becomes increasingly challenging with multispectral data which can quickly amass hundreds of files in large file sizes for a single project. For example, just 14 imaged pages of a manuscript resulted in 940 files between the raw and processed images and totaled 89GB of data. Management begins at the most basic level with file names. PRTD established an unwavering file naming protocol that not only directly indicates the object's catalog number but provides information about the contents of the image like its resolution, the page number of the object, etc. which allows for a quick understanding of the image's composition without having to open the file to ascertain this information. PRTD also uses a strict folder name hierarchy for separating relevant content that all their multispectral projects and other analytical lab instrumentation adhere to. This congruency enables multiple users with varying levels of experience to easily access, navigate, and understand the data even years later. PRTD has engaged in a larger scale initiative called The Center for Linked Analytical Scientific Samples - Digital (CLASS-D) to link a range of scientific instrumentational analyses back to the original cultural object or reference material and further the capability for access of this data [4].

Data storage often poses a substantial obstacle for multispectral imaging data sets in the ability to accommodate the required amount of space and diverse content. Not only does the data need to be securely archived, but a working copy must be retained for access by staff and researchers. As described above, all data must be in lossless file format. These particular workflows depend on and ultimately require the implementation of repeatable and audited processes to ensure verification of data integrity. The Library of Congress uses a network Content Transfer System (CTS) which utilizes a file packaging format ideal for digital content kept as a collection of files. This process employs a checksum algorithm for verification with corresponding metadata to document storage and transfer and to ensure repeatability.

Analytical Capabilities

The Library of Congress has been motivated to expand the use of multispectral imaging for scientific analysis beyond that of its stereotypical use. While it is an excellent means of uncovering faded or redacted information, its potential goes much further. PRTD has encouraged curators, conservators, and scholars to ask complex analytical questions rather than assume the limitations of the instrument. Developing a means to answer these questions became a driving motivation to push the system's capabilities. PRTD's multispectral system uses a compilation of diverse illumination modes within a single capture sequence to extract all possible visible and non-visible information from the object into a comprehensive image cube. Often a combination of these illumination modes is needed to answer analytical questions about a collection item, but in some cases, all are used to create a thoroughly comprehensive baseline characterization of an object with uncertain origins.

The sequence begins with standard reflectance imaging in each available wavelength. This provides the core set of data which distinguishes all the various materials (substrate, inks, pigments, stains, residues, etc.) that comprise the object by discerning their unique spectral characterizations across the UV-VIS-IR spectrum.

Low angle raking illumination in each end of the spectrum (470nm blue and 910nm infrared) skims across an object and provides a topographic mapping of the surface. This provides information about historic construction techniques by enhancing the impressions of woodblock imprints, compass holes, parchment manufacture, and even indented writing.

A filter wheel containing 6 broad band filters in various combinations with ultra-violet (365nm) and deep blue (450nm) light is engaged to augment fluorescence emission, ultraviolet reflectance, and ultraviolet absorption which not only enhances non-visible materials but can provide identification of their composition based on these unique characterizations. Imaging under ultraviolet light can be helpful in identifying pigments, adhesives or other residue, and evidence of cleaning or conservation treatment.

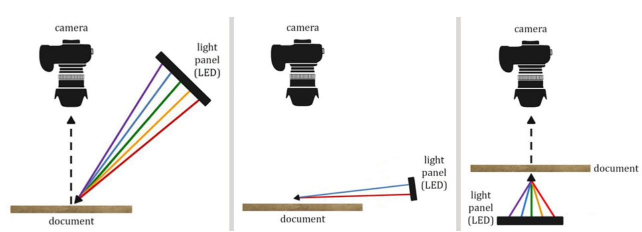

Figure 4. PRTD's standard imaging modalities for baseline characterization: reflectance (left), low-angle raking (middle), and transmitted (right).

Multispectral transmitted illumination, which projects light through an object from beneath it, is used to identify watermarks and establish provenance of an object, identify inconsistencies in the substrate material, or to look through layers like paste downs or laminations. The use of *multispectral* transmitted illumination is specifically important as the spectral variation of materials enable the digital "removal" of interfering components. For example, when capturing the watermark in a paper used for writing a letter, transmitted white light would render the watermark obscured by the written text. However, with multispectral transmitted illumination, the written text could be dropped out through image processing, revealing a clear image of the watermark.

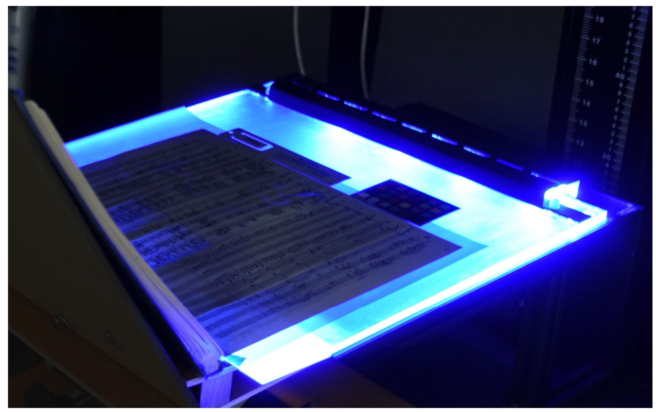

Figure 5. Multispectral transilluminator wedge during capture of a bound manuscript

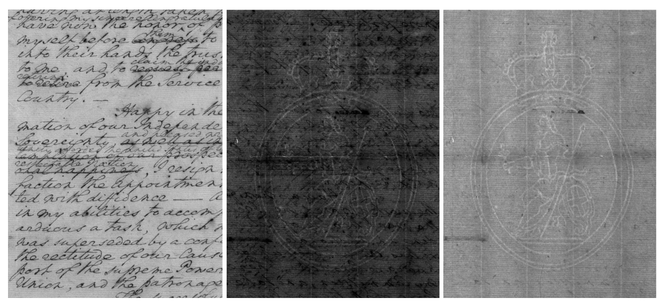

Figure 6. Letter containing a watermark in the paper, as seen in visible reflected light (left), visible transmitted light (middle), and as a result of multispectral variation of the ink to allow drop out of text (right)

Multispectral imaging is known for its ability to enhance faded and obscured text like that found in palimpsests, redacted documents, erasures, or objects that have suffered damage or staining. This is primarily accomplished through image processing procedures like principal component analysis (PCA), a mathematical algorithm that detects and highlights differences in spectral responses between the various material components of an object. PRTD utilizes these procedures but also expanded the use of PCA to scrutinize variation within a single component, emphasizing differences in the pressure of an ink's application on paper to create a unique identifier of authorship in instances where documents were scribed in a similar handwriting style.

Spectral curve analysis is enabled though the standardization of workflow and careful documentation of the equipment. A spectral curve is a visual representation of a medium's spectral data that illustrates the reflectance of light from its surface (y-axis) wavelength-by-wavelength (x-axis) forming a unique curve. This allows for qualitative comparison between the similarity of materials and for the identification of media, especially inks and pigments, by comparison to known references. Spectral curve analysis is also an effective means of analyzing change over time. This is useful when tracking an object on exhibition display, the impact of conservation treatment, and the artificial aging of scientific reference samples. The careful calibration and routine measurement of the imaging equipment, as discussed earlier, eliminates variables of error to ensure precise and accurate comparison between multiple imaging sessions so that materials do not need to be in the same imaging scene to be accurately compared.

PRTD has found success in adapting non-standard accessories to expand the range of materials the sole instrument can capture. It was through this type of research and development that multispectral transmitted illumination became a standard mode of capture in PRTD's workflow. The use of anti-Newton Ring glass enabled unobscured imaging of transparent media like film, which otherwise contained distortion in the captured image. The incorporation of a 400nm LED, which sits on the border of the UV and deep blue spectrum, facilitated extraction of data comparable to UV induced information (fluorescence, residue, etc.) through UV protective glass of an object's encasement which was blocked at 365nm.

Multispectral imaging is an efficient technique to quickly map an entire object and differentiate its material composition. It makes for an excellent preliminary analysis to complimentary point-source analyses like X-ray fluorescence (XRF), Raman spectroscopy, and fiber optic reflectance spectroscopy (FORS). For example, two inks that visibly look the same but are actually of different chemical composition are easily discerned with multispectral imaging but with XRF alone may be overlooked and only one of them analyzed. Multispectral imaging can additionally support these complimentary analyses by providing a means of confirming analysis results like pigment and ink identification.

Integration

The integration of a multispectral imaging system and its implementation for scientific analysis may seem challenging as it impacts an institution's staff, existing workflows, and data management systems. It requires adaptation and the balance of practical application with available resources and time without sacrificing quality, reinforcing the need for and adherence to standards and guidelines. These requirements include:

1. Collaboration with curatorial or conservation staff for the procurement, preparation, and handling of the object.
2. Specialized training of system operators for image capture and subsequent quality control procedures of the collected data.
3. Advanced training in image processing for the analysis of spectral data.
4. Data management of large file sets and corresponding metadata to meet established institutional standards and ensure data integrity.
5. Research and collaboration with curators and scholars to provide historical context and generate considerations for additional approaches of analysis through their familiarity with the object, in turn ensuring processed images accurately reflect the features of the object.
6. The capacity for storage, archival, and access of these comprehensive scientific datasets.
[5]

Conclusion

Expanding the use of multispectral imaging for documentation, scientific analysis, and non-invasive characterization of materials helps to advance the preservation of cultural heritage collections. Its wide range of applications include the characterization of inks, colorants, substrates, treatments, assessment of environmental factors, and recovery of obscured or deteriorated document content information. The development of a standardized procedure that is repeatable and reproducible allows the accessibility of this technology in a credible scientific approach.

References

[1] FGDC-STD-001-1998, "Content standard for digital geospatial metadata".

[2] T. Rieger (Ed.) "Technical Guidelines for Digitizing Cultural Heritage Materials", Federal Agencies Digital Guidelines Initiative, 2016.

[3] ISO 15836-1:2017, "Information and documentation – The Dublin Core metadata element set – Part 1: Core elements".

[4] F. France, "Advances in Integrated Research Infrastructures for Science and Humanities Linked Data", Imaging Science and Technology, Archiving 2017 Final Program and Proceedings, pp. 39-43(5), 2017.

[5] F. G. France, D. Emery, and M. B. Toth, "The Convergence of Information Technology, Data and Management in a Library Imaging Program", Library Quarterly special edition: Digital Convergence: Libraries, Archives, and Museums in the Information Age, vol. 80, No. 1, pp. 33-59, 2010.

Author Biography

Meghan Wilson, a Preservation Science Specialist in the Preservation Research and Testing Division at the Library of Congress, has worked extensively with multispectral imaging technology, developing guidelines and workflows for technical operation of equipment and image quality control of capture and processing procedures. She specializes in advanced image processing techniques and is experienced in other non-invasive complimentary analyses. She has applied these skills to support cultural heritage institutions worldwide in the operation, training, quality control, documentation, and data management of spectral imaging systems.

Dr. France, Chief of the Preservation Research and Testing Division at the Library of Congress, researches spectral imaging techniques and addresses integration and access between scientific and scholarly data. An international specialist on environmental deterioration to cultural objects, her focus is connecting mechanical, chemical and optical properties from the impact of environment and treatments. Serving on standards and professional committees for cultural heritage she maintains collaborations with colleagues from academic, cultural, forensic and federal institutions.

Chris Bolser, a Preservation Technician in the Preservation Research and Testing Division at the Library of Congress, applies his forensics background to the analysis of cultural heritage materials. He is proficient in an array of advanced image processing techniques and has contributed to the development of standards and procedures of preservation imaging analyses.