DOI: 10.2352/issn.2168-3204.2016.1.0.6

Burned Record Imaging at NARA

Noah Durham; National Archives and Records Administration; St. Louis, Missouri, USA

Abstract

On July 12, 1973 a devastating fire at the National Personnel Records Center (NPRC) in St. Louis destroyed approximately 16-18 million Army and Air Force Official Military Personnel Files (OMPFs) from the early and mid-twentieth century – 80% of the total military personnel records of this time period. During the salvage efforts following the fire approximately 6.5 million records were recovered. OMPFs are commonly requested by government agencies and individual citizens for the purpose of obtaining veterans' benefits and researching personal or family histories.

In 2015 a multiple spectrum imaging workflow was put in place to digitize the most heavily damaged records as they are requested – records at a condition level that formerly prevented their being accessed by requestors. Digital versions are provided in lieu of originals for these requests. To counter information loss in charred record pages, the National Archives and Records Administration (NARA) utilizes infrared imaging as part of the process.

This paper describes the production imaging project now underway, with particular attention paid to the unique requirements of photographing brittle and charred originals. The author shares information about the overall strategy, methods to prepare records for imaging and a photographic technique to maximize content recovery.

Access to the Fire-Affected Records

The National Personnel Records Center (NPRC) is the NARA repository of personnel-related records for both the military and civil services of the United States Government. It's among NARA's largest operations with 4 million cubic feet of record storage and over five hundred employees. Customers of the Center include other government agencies, military veterans and their family members, former civilian Federal employees, and the public at-large.

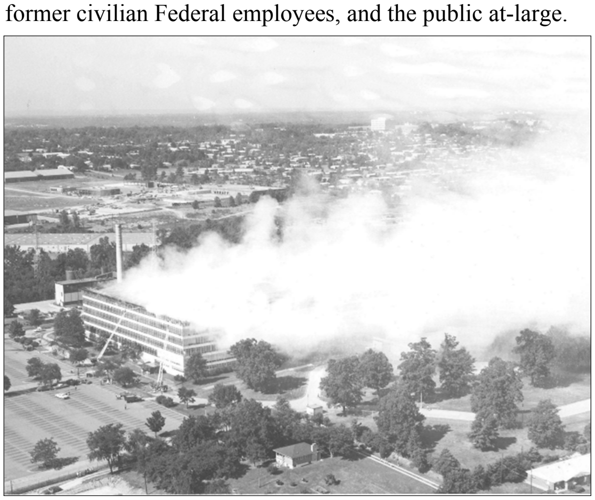


Figure 1: *Scene from the 1973 Fire. After the fire NARA continued to occupy the same building (with the top floor removed) until relocating to a new facility in 2011; fire-affected records are stored in temperature and humidity controlled bays separate from other holdings.*

Fire-affected Military Personnel Records, referred to as "B-files", or "burned" files, by NARA staff, are some of the most heavily demanded records within the NARA system, with approximately 200 record requests occurring daily from among the 6.5 million salvaged records.

The NARA Preservation Branch in St. Louis supports reference operations by cleaning, repairing and stabilizing B-files so they can be safely handled or photocopied by staff or a requesting agency or used by the public within the public research room. Preservation Technicians inspect all records prior to release. The extent of existing damage varies considerably and records may contain inactive mold spores which could be unsafe to staff and public if not removed. Fortunately most records are serviceable and are released after some combination of mold inspection and remediation, surface cleaning, flattening, and/or mending.

Figure 2: *A highly damaged B-File showing a full range of charring from light browning to highly brittle fragmented char.*

Description of the Project

NARA's goal is to enable access to every damaged record requested, or at a minimum the information that the record provides. However, in the case of very deteriorated or brittle records that would become further damaged if handled, or if a record can't be read because the content is heavily obscured or pages are fused, the records are withheld. In October 2015, the St. Louis Preservation Branch began to digitize the most heavily damaged B-files (e.g. Figure 2) to enable access to them as they are requested.

Records are prepared by Preservation Technicians who carefully remove fasteners (staples), perform any necessary surface cleaning and transfer pages into polyester film sleeves using a micro-spatula. The order of pages is preserved as found in B-file folders, and any loose fragments that can't be associated with a parent page are combined into polyester film sleeves at the end of the record. Once stabilized in this manner, documents can be safely digitized by a Preservation Technician with relative speed and minimal risk using an overhead camera system.

DOI: 10.2352/issn.2168-3204.2016.1.0.6

Figure 3: *A good candidate for digitization, prior to its placement in polyester sleeves.*

Figure 4: *A stabilized camera-ready b-file can be fully imaged in both visible and infrared in less than a minute per page, cropping included.*

Two images are normally captured for each document page; one conventional color image and one infrared image. The purpose of the infrared image is to convey content in charred areas. Since beginning the project, in October 2015, 136 digital records containing 25,760 document images have been produced. Importantly, the efficiency of the project is on par with records reformatting and preservation treatment actions for other archival records. The St. Louis Preservation Branch estimates spending no more than a minute per page to fully capture even the most challenging B-file page in both conventional color and content recovery infrared images.

Requestors receive high quality access versions in multi-page PDF format. The digital version then becomes the replacement for current and future reference purposes. NARA retains a higher resolution master for each page (approximately 100MB per image). Preservation benefits are attained through the project because original records are returned to storage without being subject to handling or photocopying outside the lab environment.

Information Recovery Effects

Museums, libraries, and archives have all utilized technical imaging to improve readability of faint, degraded, obscured or altered text in their high-value holdings. Imaging and conservation scientists have collaborated with subject matter experts from a diverse range of professions including art, archeology and history to make new and exciting discoveries about culturally significant items. Given the volume of pages NARA is digitizing upon request in this project, the information recovery method it employs must be predictably successful and capable of automation for fast throughput.

Infrared film photography is the earliest successful method to dramatically improve legibility in degraded or faint text; its use was described by Kodak scientist Walter Clark in 1939[1]. Burn damage infrared imaging procedures for the aim of recovering content in forensic document exams were published in 1953[2].

NARA's St. Louis Preservation Branch evaluated both infrared and visible image captures to determine the potential of each for automated content recovery in charred B-Files. With the common Kodak 87C and 89B pass band filters, printing ink and typewriter ribbon ink noticeably contrasted with surrounding char regions in infrared captures (see Figure 7). This recovery effect was repeatable (successful from page-to-page) and could be attained without localized image manipulations.

By comparison, visible image captures exhibit faint features in printed ink, typewriter ink or handwriting where content is present in charred areas. They benefit from image processing adjustments to increase contrast in local areas to improve legibility.

Since B-files range between 50 and 200 pages per record, manual enhancements like Photoshop curves and levels adjustments are not practical or economic. The image processing technique with the greatest potential for automation with a predictable result is a channel calculation. Due to differences in red and green absorption for ink and char features in light to moderate charring, the difference in values R and G channels can be computed and represented as a new channel for readability. This technique was described by Knox, et al., in the discovery of hidden content in the Dead Sea Scrolls [3]. The author does not have a RGB conversion to test the same color space as Knox used, but visualizes this effect (Red minus Green) by using Photoshop Channel Mixer in Figure 8.

However, NARA's tests of visible spectrum captures show an upper limit on content recovery potential. When the charring is extreme, minimally visible content features diminish and disappear altogether; so also does their recoverability. Figure 6 shows the zone where only infrared capture, not visible, could recover content. As a result, NARA selected infrared capture as the better starting point for content recovery.

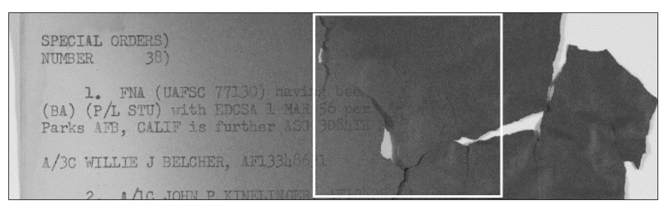


Figure 5: *This standard color image is highlighting the moderate char zone where some visible content is minimally legible and could be image processed to improve readability.*

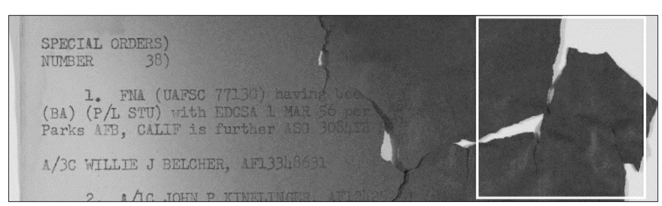
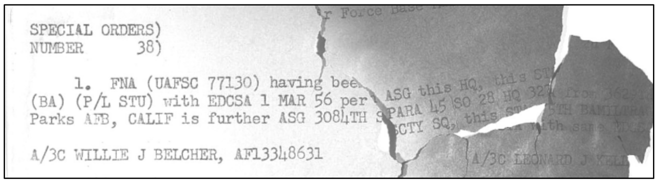


Figure 6: *This standard color image highlights the heavily char zone where no visible content remains and only infrared was able to recover content.*

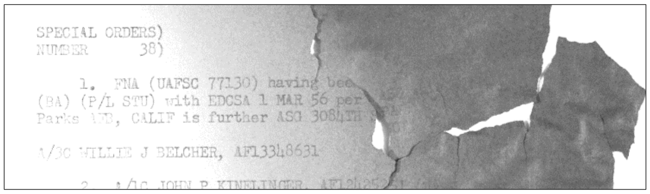
DOI: 10.2352/issn.2168-3204.2016.1.0.6


SPECIAL ORDERS)
NUMBER 38)

1. FNA (UAFSC 77130) having bee
(BA) (P/L STU) with EDCSA 1 MAR 56 per
Parks AFB, CALIF is further ASG 3084TH

A/3C WILLIE J BELCHER, AF13348631

Figure 7: *This is an infrared image showing text recovered in even in the most extreme charred area*


SPECIAL ORDERS)
NUMBER 38)

1. FNA (UAFSC 77130) having bee
(BA) (P/L STU) with EDCSA 1 MAR 56 per
Parks AFB, CALIF is further ASG 3084TH

A/3C WILLIE J BELCHER, AF13348631

Figure 8: *This is a visible capture, where Photoshop Channel Mixer was used to amplify the difference in Red and Green channels across the charred areas and enhance readability in light to moderate charring.*

Camera Selection

The DT Multispectra camera incorporates a PhaseOne digital back with a 60 megapixel sensor. Its high pixel count (8984 x 6732) is more than adequate resolution for conventional color imaging of records, but is every bit needed for infrared imaging; where acquiring the clearest possible signal boosts the content recovery results. The DT camera can be configured with PhaseOne backs of 40, 60, 80 or 100 megapixel resolution. For the shooting height that NARA uses, the 60 megapixel capture results in 690ppi resolution for actual size reproductions.

Prior to purchasing the DT Multispectra (~$59,000 complete with lighting and computers) NARA also tested two other camera types with lower resolution and smaller sensor sizes: the Fuji IS Pro forensic camera (~$1200) and an infrared modified Canon 5D Mark II DSLR (~$5,000). These experiments gave NARA the opportunity to evaluate infrared recovery effects at lower resolutions. Like the DT Multispectra, these cameras have internal infrared cut-filters removed, permitting infrared wavelengths to reach the camera sensor. Technicians place filters in front of the lens to select an infrared or visible shot type.

The Fuji limitations are what one expects for compact cameras with a small sensor: noise levels are high, especially in infrared captures. Despite having quality limitations, the Fuji performed well in focusing and exposure even with an optically dense infrared filter in place. While not well suited for this application, it is useful as a portable hand held camera in both visible and infrared modes. For a conservation lab the Fuji would be an ideal choice as a bench tool for instant infrared analysis of an item. A magnetic filter holder can be used to facilitate easy filter changes.

The infrared modified Canon DSLR is a higher resolution and dimensionally larger sensor (20 megapixel full frame DSLR). As expected, the quality of the Canon infrared capture was superior to the Fuji. The Canon can be operated in tethered mode (connected to a computer) while on a copy stand, making it easier to transfer and access files in editing programs. The Canon is more sophisticated than the Fuji, and paradoxically this was a problem. Auto-focus, live view and light metering often failed when the unexpectedly dense IR filter was introduced in front of the lens. It's likely that the Canon can be configured to work more predictably in future tests, but it seemed highly tuned for pictorial photography, rather than technical work.

Figure 9: *The DT Multispectra camera with a 60 megapixel camera back. Focus locks and filter turret make it ideal for adapting shot types in consecutive shots.*

Photographic Methods

The project is on pace to capture between 70,000 and 100,000 images annually.

NARA operates the camera like a book scanner; meaning the head height is kept fixed to accommodate the largest size original we encounter in B-files (11"x14"). Similarly, lighting position, f-stop and shutter speed are set and not changed by technicians throughout the camera session.

An imaging session begins by setting the focus stops for both visible and infrared images. Technicians test and adjust exposure by reading RGB values from a photographed test target. Once focus and exposure is set, the camera is simple to operate in production. Technicians alternate between visible and infrared shot types, selecting the filter and focus setting to match.

Because char – independent from polyester film, has a fairly glossy surface (see Figure 2), recovery effects can be hindered by highly directional light creating shadowing and specular highlights within the image. The position (angle) and amount of diffusion in lighting are important considerations for image quality. With infrared images, diffused lighting from conventional copy angles (like 45 degrees) worked best. However in conventional color images, glare from the surface of polyester film will be apparent if lighting is too soft (Figure 12). Accordingly, a balance needs to be struck to find the lighting angle and extent of diffusion that works for both shot types.

DOI: 10.2352/issn.2168-3204.2016.1.0.6

Waviness (planar distortion up to 1/16 of an inch) is inherent to charred paper. The weight of the polyester film, 3 mil (.003 inch), does not aggressively flatten the waviness, instead it constrains the document and improves flatness (see Figure 10). Contrary to what might be expected, flattening, even gently, with a book cradle or glass platen creates additional waves and lowers the quality of images. Flattening is also a risk for further fragmentation of these very brittle sheets, therefore B-file pages are laid directly on a copy board during image capture.

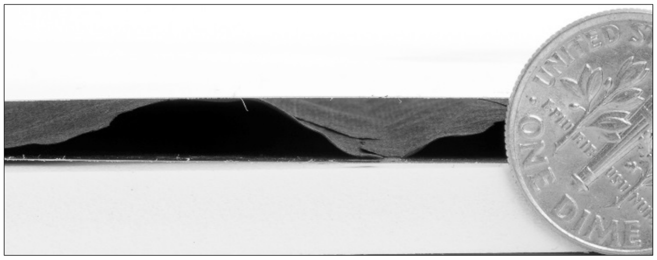


Figure 10: *A side view image of a charred paper in polyester film. The surface wave is contained but not fully flattened. Dime provided for scale.*

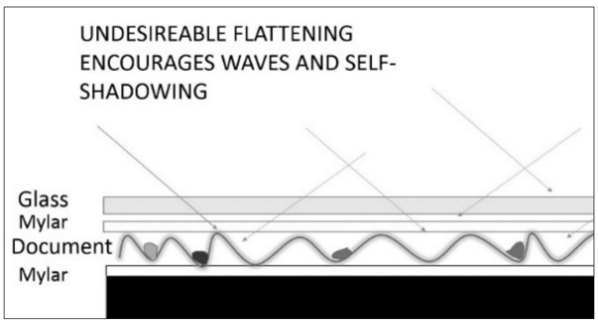


Figure 11: *Additional waves result from flattening between polyester (Mylar®) and glass.*

The most practical light source for production that NARA tested was strobe. Strobe gives an acceptable balance of light quality and strength in both the visible and infrared capture. However, there is a four (4) stop difference in optimal exposure between visible and IR shot types when using strobe lighting. A software compensation of four (4) stops is performed to brighten infrared exposures prior to export. Although the dynamic range of the sensor can accommodate the compensation and it can be performed automatically, this compensation adds noise to the image, decreasing quality.

Tungsten lighting, having a high amount of infrared output, performed better than strobe in testing. It had superior image quality and did not require exposure compensation between shot types. However, in real world production it was impractical. Tungsten emits too much heat; during long imaging sessions lasting up to a full day the hot lights become an ergonomic and safety problem. Therefore tungsten lighting was ruled out. In the next section *Future Research and Process Improvements,* potential solutions utilizing a mixed grid of LED light panel is discussed.

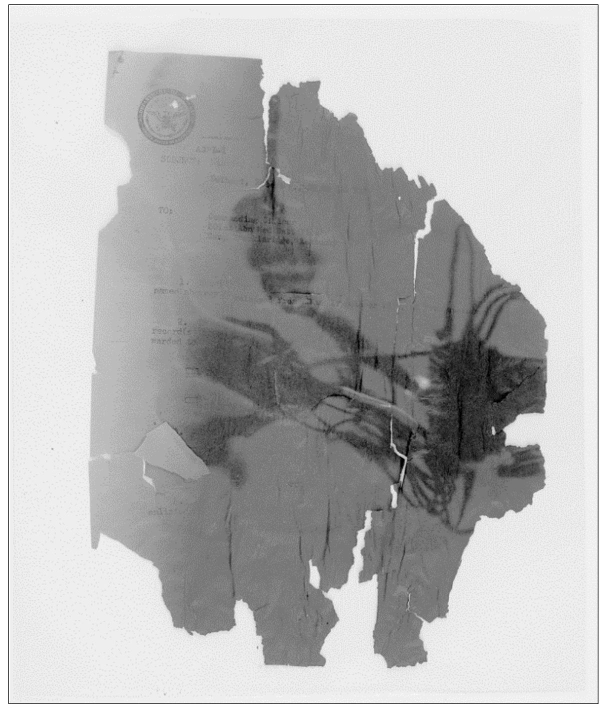


Figure 12: *An example of a color image. Poor lighting causing glare from the polyester film surface.*

Figure 13: *The same document in Figure 12, but infrared image. Tones that define the surface features compete with text readability. The best lighting is soft enough to evenly light the surface, but not so soft as to create glare on polyester film surface.*

DOI: 10.2352/issn.2168-3204.2016.1.0.6

Future Research and Process Improvements

Improvements continue to be made to the overall project. The Preservation Branch is conducting trials with wet chemical treatments to separate fused document pages to provide access to B-files whose sheets are fused together. It will be interesting to see how paper distortion (waviness) is affected by these treatments and if there is any impact on image quality. Generally, when documents lay flatter, there are fewer surface features to interfere with infrared readability.

Work continues on narrowing if not eliminating the exposure compensation between shot types. Preservation is testing a hybrid approach of strobe lighting with supplemental infrared fill lighting. Such a strategy would improve image quality of infrared image, by reducing noise, while not impacting the conventional color image in the least. The Preservation Branch is also considering a custom LED light panel as a source, instead of strobe. With an array of both visible and IR diodes, the percentage of each diode type could be optimized to balance the visible and infrared output, to maximize the recovery effects, and eliminate the need for software exposure compensation.

NARA's Preservation Branch hopes to collaborate with a scientific lab, or other institution with the ability to spectrally characterize the features within charred records. Although the inks, paper and time period cover a broad span in B-files, it's conceivable that a single or a few effective wavelengths enable a strong precise recovery effect. It is also possible that spectroscopy could point to a different pass band filter or frequency range that's key to imaging certain recurring forms, such as military induction and separation forms, more effectively.

Lastly, the master files NARA captures contain a lot more information (690ppi) than what is delivered to requestors in lower resolution access versions (300ppi). The Preservation Branch would like to set up a web portal with deep zoom capability that allows selective zoom and resolution, to economically serve out higher resolution "regions of interest" as users zoom in.

References

[1] Walter Clark, Photography by infrared; its principles and applications, J. Wiley and Sons, 1939

[2] Donald Doud, "Charred Documents, Their Handling and Decipherment," Journal of Criminal Law and Criminology, 1953

[3] K.T. Knox, R.H. Johnston, and R.L. Easton, "Imaging the Dead Sea Scrolls," Optics and Photonics News, Vol. 8, No. 8: 31-34, August 1997.

Author Biography

Noah Durham is a Supervisory Preservation Specialist for NARA. His professional interests stem from an early fascination with photography, photogravure and other printing systems. Noah holds an MS in Graphic Arts Systems from the School of Printing Management and Science at RIT. In the 1990s he set up the first large scale digital photography operations for the auction houses Christie's and Sotheby's, in New York City, London and Amsterdam. Throughout his career Noah has been leading efforts to solve complex problems in pre-press, printing, imaging and archival operations. For the last 8 years Noah has worked for the National Archives in St. Louis, Mo, where he supervises a preservation reformatting lab for the National Archives managing large scale preservation projects for both paper and microfilmed records.

