Picking up the Thread: Implementing a New Digitization Quality Monitoring Program at the National Archives

Nabil Kashyap

Introduction

New tools and shifting priorities have made large-scale image digitization projects commonplace; however, most documented best practices and suggested standards do little to assess achieved image quality. Guidelines geared specifically toward image quality in digitizing heritage materials, such as Metamorfoze and FADGI Still Image Working Group Guidelines, have recently emerged in part to fill that need -- yet how do sponsoring institutions actually implement these technical metrics and stringent standards given the scale and complexity of their digitization projects?

The Photographic Imaging Lab (PI), part of the National Archives Digitization Services Branch, is currently transitioning to a quality control process based on FADGI Guidelines star-rating system. PI currently relies on a custom, hybrid grayscale/color separation imaging target derived from the Kodak Q13, originally created to monitor photographic processes. The lab intends to adopt Image Science Associates GoldenThread system of integrated targets and analysis software -- but for each new potential metric, new decisions must be made about appropriate standards, what data to collect and how to monitor that data.

Methods

Taking into account available resources, organizational concerns and current practices at PI, we drew up a comprehensive proposal for implementing a new quality monitoring program designed around GoldenThread. The proposal is based on research of existing literature, firsthand observations, user and expert interviews and hands-on testing that took place over a nine-week period.

We made direct observations of all phases of production across most pieces of equipment currently in use. We conducted both formal and informal interviews with users from across the division, including users from PI, from partner labs and from administration. In addition we consulted with imaging specialists Steven Puglia, Library of Congress, and Don Williams, designer of GoldenThread targets and software. Hands-on testing included spectrophotometric measurements of targets, analyses across multiple versions of GoldenThread software and individual inspection of image files. Lastly, we processed sample scanning jobs over several devices in order to establish baseline performance and to test a pilot workflow.

Goals

To ensure products meet image quality guidelines

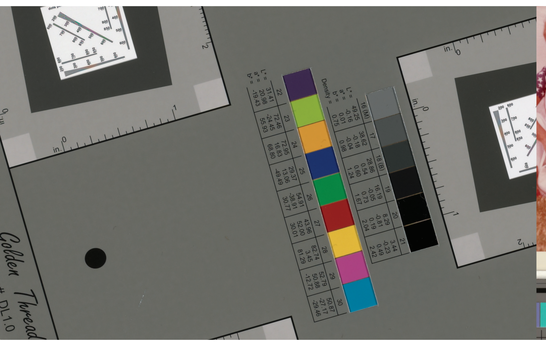
Currently the lab cannot assess to what degree products meet FADGI guidelines, especially with regards to tonal response, color-encoding accuracy and achieved resolution. First, current item-level targets do not contain the features necessary to test achieved resolution and vary too widely due to printing inconsistency and wear to test tonal response and color-encoding accuracy. Second, workflows vary somewhat across devices and jobs; information obtained through current targets is interpreted differently by imaging technicians and quality control technicians depending on context, which prevents products from being meaningfully compared across jobs.

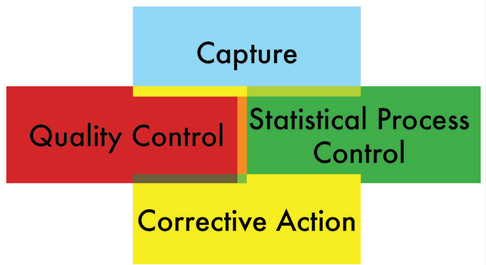
To ensure capture devices are capable of products that meet image quality guidelines

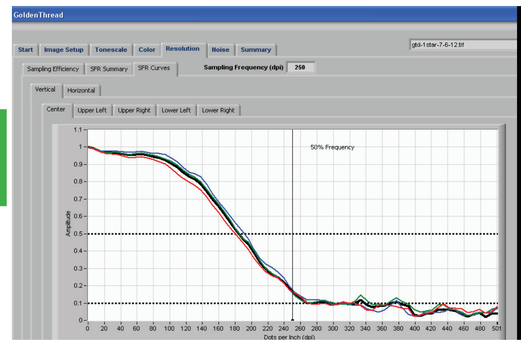
Maintenance at PI is largely oriented toward catastrophic failure. PI has no mechanism in place to systematically assess device performance and compare performance across devices. While some PI devices, like Zeutschel scanners, have device-specific targets and analysis software provided by the manufacturer, most devices in use do not, which leads to inconsistency in device maintenance.

To reduce the need for recapture due to image quality issues

PI regularly rescans items during larger jobs due to quality control issues. This duplication of effort is a clear drain on resources. Systematically monitoring device performance and taking advantage of more precise targets during initial capture could preempt certain quality control issues and, by avoiding some reprocessing, improve efficiency.







Underpinning our proposal are four interrelated layers: capture, quality control, statistical process control and corrective action.

Capture

The capture layer gathers information about individual captures and about equipment performance with only minor additions to current practices, scanning device-level Golden Thread targets before and after every session and replacing current item-level targets with Golden Thread targets. In light of the number of technicians, the variety of devices and the variety of job specifications, a systematic, reproducible workflow is critical to collecting reliable, actionable data.

Quality Control

The quality control layer inspects products, assessing whether image files meet specifications, either passing files along for delivery or returning files to technicians for reprocessing. This involves batch processing a random sample of image files, the size of the sample either being all files in a job or a statistically appropriate proportion of the total job.

Statistical Process Control

The statistical process control layer (SPC) monitors device performance over time, assessing when a device operates according to specifications, operates at a warning level and operates at a level unacceptable to lab aims. Though SPC does not yet formally exist at PI, its aims and methods are closely aligned with the current quality assurance procedures informally in place. Where quality control is largely a binary process, using defined criteria to pass or fail files, SPC tracks device performance in order to empirically determine when to perform preventative maintenance.

Corrective Action

The corrective action layer responds to data interpretation by identifying causes of device performance issues and attempting to resolve them. No amount of data collection or analysis is ultimately useful unless there are processes in place for acting on what that data tell us. A corrective action layer consists primarily of two parts. First is a set of documented procedures across capture devices that respond to specific image quality issues. Second is a system for tracking and scheduling maintenance performed through support contracts and performed in-house by staff.

Conclusions

We found that an effective quality monitoring program should be tuned to the capabilities of an individual lab. Different devices and capture environments affect imaging uniquely, and in order to make best use of quality data, a lab needs to establish baseline performance profiles. There is no literature yet about how often devices need to be monitored, but based on input from imaging specialists, we recommend scanning the device-level target before and after each capture session for an initial period of ninety days in order to create adequate device profiles.

One current underlying the four layers described above is widespread imaging literacy among users. We find the more staff that are able to interpret quality measurements, the more effective a quality monitoring program. Imaging involves an extraordinary range of variables, and a realistic quality monitoring program cannot operate as an automated pass/fail gate. In order to ensure a quality monitoring program supports rather than confounds larger project aims, we see image literacy among technicians, supervisors and administrators as a key component.

We see these layers as being applicable to any large digitization project, regardless of institutional affiliation or of available resources. While this proposal is structured around a specific target and software system, GoldenThread is not the only integrated target nor is it the only image quality analysis software available. New options, including printing targets in-house and analyzing targets using free, web-based applications, now make image quality standardization an option for a range of projects.

An important question beyond the scope of this proposal is how to coordinate imaging standards when different departments within an institution sponsor significantly different digitization programs. While PI has invested resources into this program, the lab does not account for all of NARA's digitization. Many institutions with vast holdings and complex organizational trees, like NARA, have compartmentalized digitization projects that have evolved organically over time -- from scanning within custodial units to public Scan-A-Thons to for-profit partners like Ancestry.com and Fold3. Ideally, an archive should be able to articulate the standards by which it produces all products, yet how we articulate those standards across the diversity of digitization projects happening today is an open question.

school of information
University of Michigan
INSTITUTE of Museum and Library SERVICES
NATIONAL ARCHIVES