When Good Scanning Goes Bad: A Case for Enabling Statistical Process Control in Image Digitizing Workflows

Michael Stelmach, Library of Congress; Washington, DC
Don Williams, Technical Advisor, Rochester, NY

Abstract

Institutions invest tremendous amounts of time, money, technology and human capital to digitize their collections. But surprisingly, few apply even basic statistical process control (SPC) strategies to monitor the output quality of their imaging workflows. Perhaps it's the misplaced notion that digital imaging is error free or losslessly correctable. Maybe hardware manufacturers have seduced users into assuming that image quality is a foregone conclusion. Or, perhaps institutions would welcome such strategies but are simply not enabled with appropriate resources and knowledge to effectively practice them. All three of these play a role, but we believe the latter is the greatest obstacle to implementing SPC strategies.

The benefits of SPC in industry are recognizable in terms of quality, efficiency, and economy. The same can occur with digital imaging in the cultural heritage community. Any SPC program naturally involves monitoring selected output parameters. The short list could be sampling rate (dpi), resolution, noise, and tonal/color fidelity. By way of a suitable target artifact and analysis software, each of these variables can be periodically measured and their values compared to pre-established numerical aims and error bounds. Corrective action is taken when trends approach or exceed these bounds. This paper describes efforts at the Library of Congress to introduce SPC practices into imaging workflows by supporting the development of unique hardware and software tools that provide ISO standardized imaging performance measurements.

In addition to advocating the incorporation of genuine SPC practices into the imaging workflows of cultural heritage institutions, we also present some encouraging progress on enabling ISO imaging performance compliant tools to accomplish this in a workflow and archiving friendly fashion. Results exercising these tools are shared, and future approaches are presented. Because of its fundamental importance, the image capture stage is the focus of this paper. However, the principles involved apply equally to display, printing, and metadata generation.

Introduction

> *Opportunity is often difficult to recognize; we usually expect it to beckon us with beepers and billboards.*
> *~William Arthur Ward*

These are both exciting and taxing times to be involved in digital conversion. A plethora of new opportunities and challenges exist today, and news of new opportunities seems to come out every week. The vast sums of money behind commercial digital photography have brought about rapid development in instant capture that is now available to our community. On the high end, devices such as the Phase One® P45 digital back provide 39 MP and 48-bit color in an instant capture device. On the lower end, a variety of digital SLRs are presenting new opportunities in speed and cost, allowing libraries to undertake projects that would have been out of their reach just a few years ago. With these opportunities, however, come smaller sensor pixels, myriad lens options, flexible lighting, and a host of variables that challenge our ability to achieve reliable image quality.

Mass digitization projects of a scale barely imagined a decade ago are taking place around the world, and the pace appears to be picking up. Libraries are now being presented with the challenge of dealing with workflows that have image files in the millions per month rather than the thousands or tens of thousands they may have been accustomed to. Libraries are also entering into partnerships and collaborations at an ever increasing rate. Field work is becoming more common; an environment where digital conversion can be far from ideal. How can consistent quality be obtained and quality processes be shared?

All of this points to a need to bring SPC from industry into the library community, and to do so sooner rather than later. At the Library of Congress (LoC), we are in the early stages of moving in that direction. Rather than undertake a single project around SPC, we are undertaking a variety of pilot projects to explore integrating existing tools, and are exploring the development of new tools and procedures. We have a number of projects underway that incorporate the use of device-level and image-level targets, JHOVE as an automated component of QC, and statistical manual QC review of image files.

Before going too far in extolling the virtues of SPC, we would like to take a moment to explain why we have chosen to use the term "Statistical **Process** Control" rather than "Statistical **Quality** Control" (SQC). After all, we are focusing on product quality represented by image quality. But as practitioners of digital conversion, we have a vested interest in quality along with efficiency, and this is where SQC fails us by being too limited in focus.

With SQC there exists an end-of-stream process providing statistics used to accept or reject a product based on an acceptable quality level (AQL). In contrast, SPC methods involve more frequent sampling with mechanisms for upstream feedback to correct production processes; preventing quality problems earlier in the workflow, thereby minimizing items that fall outside the AQL. In practice, a workflow requires both what is considered classical SPC as well as SQC methods. For the purposes of this paper, we will consider SPC as inclusive of SQC.

The Manufacturing Model

> *…."The story here is not that they made a mistake in the scanning and scoring but that they seem to have no fail-safe to alert them directly and immediately of a mistake," said Marilee Jones, dean of admissions at the Massachusetts Institute of Technology. "To depend on test-takers who challenge the scores to learn about system failure is not good."…*
>
> *~ "SAT Errors Raise New Qualms About Testing" NYTimes, March 10,2006*

The manufacturing sector has methodically practiced SPC for more than two decades [1]. The reasons are twofold; due diligence and quality. Due diligence by way of best practices serves the legal and standards compliance issues, e.g., ISO 9000. But the real benefit is in terms of quality. For this community a quality product translates to durability, reputation, and longevity.

While the parallel to manufacturing may not seem apt, it is undeniable that given the high volume of digitized objects this community delivers, the analogy is appropriate. We manufacture a great number of images – high-end images that are intended to endure. In many ways we are fiduciaries for cultural heritage imagery. This is why it is so important to monitor image quality. Unlike the world of film imaging where one could confidently rely on the history-rich reputation of a few manufacturers for performance integrity, today's digital imaging landscape offers fewer assurances. The torch has been passed…to us.

As in any manufacturing process performance, issues of all sorts – some acute others glacially progressive, also occur in digital imaging. Many of these are deterministic in nature and can be identified or avoided with comprehensive hardware benchmarking and calibration at the onset. Others are random and more insidious, especially in workflows where human intervention is required. It is in this second vein where SPC can provide the greatest impact. By monitoring imaging performance with periodic tests (even image-by-image) one can identify potential trends early on, preventing not only rework but also embarrassment.

The key to doing this effectively is with good SPC practices backed by solid data; i.e., with sound scientific data based on accredited standards [2]. The data is the foundation while the practices are the execution. Murray [3] has provided an excellent review of how the former complements the latter with a concentration on the methodologies of SPC. Some excellent examples are also provided. The collected data from these examples was enabled by two critical tools: imaging targets and analysis software. These tools were designed specifically with each other and the cultural heritage community in mind and are described next. We will start with the targets' description and the rational for their design.

Targets

> *If you aim at nothing, you'll hit it every time*
>
> *~ Anonymous*

In consumer photography it has long been recognized that any *a priori* knowledge about the scene makes image quality management easier. From time to time there have even been attempts to discreetly embed targets into the frame margins of film images or exploit the unused active pixels of digital cameras to provide this knowledge. Unlike consumer imaging , cultural heritage institutions are in a better position to embed test targets directly into a scene for monitoring multiple imaging performance parameters. The short list of these include sampling rate, resolution, color, grayscale, gamma, noise, and color misregistration.

Although imaging targets are used in the cultural heritage community, they have tended to be limited in scope, clumsy, and undocumented. They offer little meaningful imaging information by which to diagnose problems or track performance – and this is critical for archiving – there is no traceability to their meaning. Think about it. Of what utility is a target if in two, five or ten years from now the physical meaning of the color or gray patches in that target or the image of that target are unknown? It would be as if one created a map and failed to label landmarks or create a legend. It would amount to just so many glyphs. A community that is second to none in the controlled digitizing of objects has largely adopted non-optimal targets for their high volume and high quality image production.

This lack of controlled imaging measures that could easily be applied by cultural heritage institutions motivated the development of a comprehensive, standards compliant and annotated imaging targets and companion software that is designed to be workflow-friendly. Developed in cooperation with selected Digital Library Federation (DLF) members and the Library of Congress (LoC), limited testing of the target has begun. This paper is intended to introduce these targets and their use in a digitization workflow to a wider community for feedback and vetting.

The importance of a suitable target artifact cannot be emphasized enough. It defines the collection ease and quality of the data harvested for SPC practices, especially if it is to be done automatically using intelligent software. Even more so, when designed properly and included as part of the repository image, the target's image can act as a link, a *Golden Thread* if you will, to the physical essence of the original object by connecting objectively traceable target values to an image's digital code values. By providing an unambiguous trail from code value to science-based standards, it is a legacy to the physical object. It is this *Golden Thread* target that we feel is a fundamental first step to realizing imaging SPC practices for the cultural heritage community.

Two levels of targets are envisioned for typical digital conversion workflows. One for device-level quality control, and the other for image-level quality control. The device-level target would fill a camera's field of view and is intended to be comprehensive and used either at setup time or at selected times when changes in imaging performance are most expected. These would include, operator transitions, driver software setting modifications, session startup/shutdown, or any camera or lighting hardware adjustments.

A great deal of information will be obtained to determine the appropriate frequency and schedule of use by tracking the variability of data values obtained during normal imaging operations.

The image-level target is less comprehensive, but is intended to be included in each and every captured image. As such it is more discreet, compact, and made to fit alongside the source object without adding significantly to the file size. By enabling 100 percent inspection, it allows for the *Jidoka* [4] principle of stopping work immediately when problems occur and preventing the production of defective items. It also serves as an authenticating branding for each and every digital image file.

Device-Level Targets

Shown in Fig. 1 is an example of a device-level *Golden Thread* target for use with a collection of 7”x 9” source documents

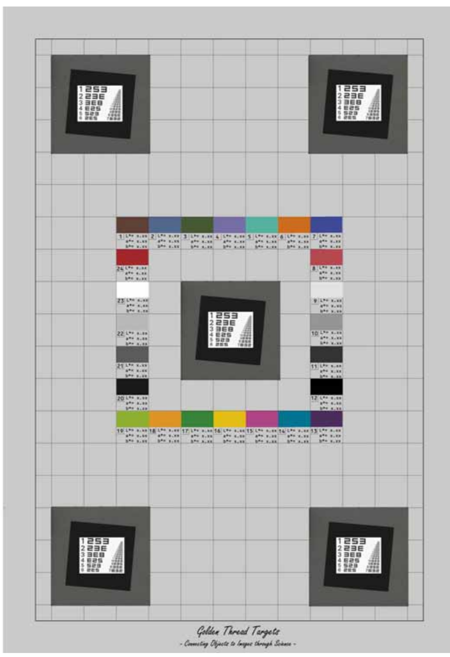

***Figure 1:** Example of full frame device-level target*

It is characterized by features that allow for the multiple and simultaneous metrology of the following imaging performance criteria across the extent of the image area. They are:

a) Sampling frequency, i.e., dots per inch
b) Resolution (true ISO resolution)
c) Visual resolution verification
d) Color and grayscale capture, i.e., gamma
e) Illumination uniformity
f) Distortion
g) Noise
h) Color channel registration error

Intended as a workflow-friendly solution for the cultural heritage community, it enables semi-automated tracking of ISO standards-based technical imaging performance. The target features are compliant with existing ISO digital imaging standards for imaging performance (ISO/TC42).

Each color and neutral patch is also documented with human and machine readable text with their CIELAB (*L*a*b**) values so future users of the digital image file can accurately reproduce the analytical colors and tones of the original object without having to search for the meaning of any particular patch. While this device-level target is not intended to be part of the finished delivered image or file, it should be part of a repository’s archived image file. Of course, input device color profiles can also be built from this information.

For consistency of past practices in this first phase, the color patch selections are from the same 18 colors and materials from the popular Macbeth® ColorChecker™. The neutral values have been expanded to 8 patches from 6 patches. The visual Status A density values of the neutrals are also documented with the target. Depending on collection characteristics, the patch colors could be changed to better reflect a collection’s predominant spectral traits.

Image-Level Targets

The next time you dispose of a boxed food package such as that for breakfast cereal, cake mix, or even dog biscuits, take a moment to break down the box by unfolding the glued tabs forming the bottom or sides of the box. Often what you'll find hidden are printer's quality control targets such as those depicted in Fig. 2. Their appearance, feature set, and utility are uncannily like the image capture targets advocated for in this section. They allow for the measurement and tracking of tone, color, resolution, and color misregistration.

***Figure 2:** Example of Printer QC targets*

Every single box that rolls off of the press has associated with it a token snapshot of the printer's performance at a moment in time that is inextricably attached to the object. This link, enables the current and future monitoring/diagnosis of imaging performance for every quantum object. How many of us can claim to take the same care in insuring quality imaging of cherished cultural objects as others do in the printing of a disposable box of consumer goods?

While such image-level targets are not as comprehensive as those for device-level testing, they can be designed to be nearly so. With prudence and ingenuity a number of similar imaging performance probes, not just standard color/tone features, can be embedded into more compact target designs. For example, Fig. 3 depicts an eight inch target in the context of the larger document. The target contains features for characterizing seven of the eight performance attributes listed in this paper under "Device-level Targets." The target is being evaluated by LoC.

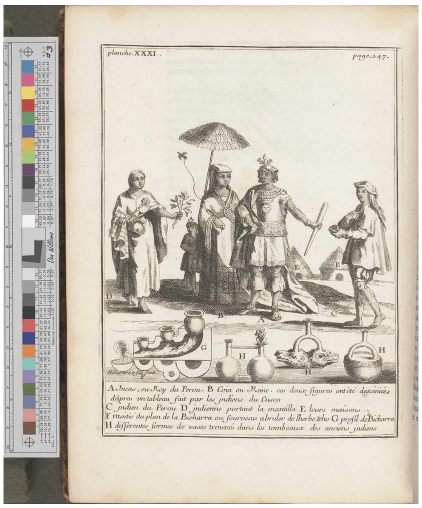

***Figure 3:** Golden Thread image-level target along side source document*
Image courtesy of Cornell University

Some maintain that adopting an image-level target strategy is overkill and unnecessary. After all, as long as periodic device-level targets are used, what can reasonably go wrong in the interim? For standard flatbed scanners where the number of imaging degrees of freedom is limited, there is a certain amount of truth to such a stance. But for more complex customized imaging devices frequently encountered in the cultural heritage community (e.g., book scanners and copy stand cameras), this can be a penny wise-pound foolish approach. Even the best of operators can get distracted or unknowingly disturb a delicate imaging setup. The dynamics of book scanning, whether operator aided or robotic, offer ample chance for mechanical movement. Plainly put, unexpected things will and always do go wrong. The aim is to catch them as early as possible on a near real time basis, automatically. That is what the proposed targets coupled with appropriate software can do.

More important, the use of an image-level target, as cited above, allows for individualized authentication; as an alternative to a batch average. It is intended to stay with the digital object much like a mattress tag; it is not to be removed, except by the end user. If the supplier feels such targets detract from the essence of the object, layered delivery files such as PDF or TIFF formats easily allow for hiding the target's image while maintaining it as part of the file.

Finally, an added benefit of image-level targets is that they free the image "manufacturer" from frequent customer inquiries on the imaging genealogy of the file. They allow the customer standardized, undigested, image-wise information that they are free to interpret and question for their own needs if desired. Is it or is it not Adobe RGB and what particular flavor? Was it truly scanned at 400 dpi or was it actually 415 dpi, as suspected? The embedded image-level target enables the verification of such scanning provenance and permits consistent, accurate, and customized image rendering.

Though imaging targets are necessary components for SPC of digital imaging they are insufficient. The complementary and equally critical portion completing the enablement is the software that, through a target's image, extracts imaging performance criteria.

Scan but Verify: Quantifying Accuracy and Variability

And now I see with eye serene
The very pulse of the machine.
~ William Wordsworth,
"She Was a Phantom of Delight"

The pulse of the machine, indeed. This is the goal –quantifiable metrics for gauging the accuracy and variability of imaging performance, preferably using vetted ISO standards. The imaged target provides the hooks. Target-specific analysis software engages the hooks yielding performance data. A block diagram of how the software might execute is shown in Fig. 4.

The role of the output data is to verify and document. As indicated earlier, Murray [2] gives several examples of how this was done using the targets and software described above. A particularly convincing example is duplicated in Fig. 5 below. It is a control chart of magnification (i.e. effective sampling rate) derived from an image-level target used in a book scanning operation. The chart in Fig. 5 clearly shows the magnification change as the scanning progresses page by page. While a two-percent error is usually acceptable, the example in shows how the magnification error gradually creeps to four percent over time.

Though we tend to think of control charts that monitor performance as a function of time, other formats are also possible. For instance, Fig. 6 was derived from device-level target data. The intent was to verify the scanner in terms of sampling rate. While most would not think to challenge the dpi settings as offered in scanner interface selections, actual sampling rate selections have frequently proven wrong compared to either the interface selected setting or the populated tagged setting. An experienced operator scanned the target at 600, 800, 1000, 1200, and 2400 dpi, as selected in the scanner's software interface. When analyzed, however, it was proven that the actual sampling rate was less than half of that selected (by a factor of 0.39 in fact). An error of this magnitude was assumed to be a simple operator oversight or software analysis bug. But upon checking the scanner interface, the operator had in fact chosen the correct setting as presented to him. The internal scanner software was plainly wrong. The suspicion is that it was actually scanning in dot-per-centimeter instead of dots-per-inch mode. This is one example of the need to verify.

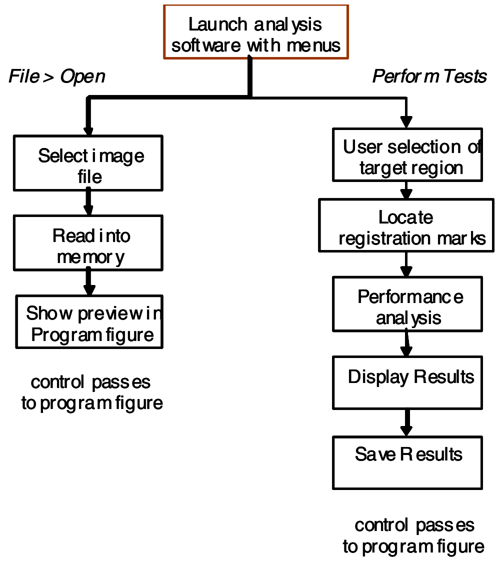

***Figure. 4:** Execution flow for the software analysis*

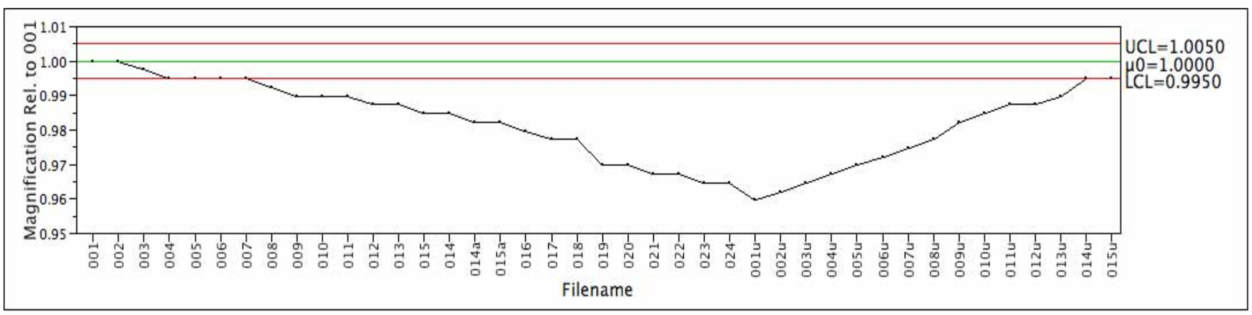

***Figure 5**: Variability in magnification for book scanning task (from Murray ref))*

Though we tend to think of control charts that monitor performance as a function of time, other formats are also possible. For instance, Fig. 6 was derived from device-level target data. The intent was to verify the scanner in terms of sampling rate. While most would not think to challenge the dpi settings as offered in scanner interface selections, actual sampling rate selections have frequently proven wrong compared to either the interface selected setting or the populated tagged setting. An experienced operator scanned the target at 600, 800, 1000, 1200, and 2400 dpi, as selected in the scanner's software interface. When analyzed, however, it was proven that the actual sampling rate was less than half of that selected (by a factor of 0.39 in fact). An error of this magnitude was assumed to be a simple operator oversight or software analysis bug. But upon checking the scanner interface, the operator had in fact chosen the correct setting as presented to him. The internal scanner software was plainly wrong. The suspicion is that it was actually scanning in dot-per-centimeter instead of dots-per-inch mode. This is one example of the need to verify.

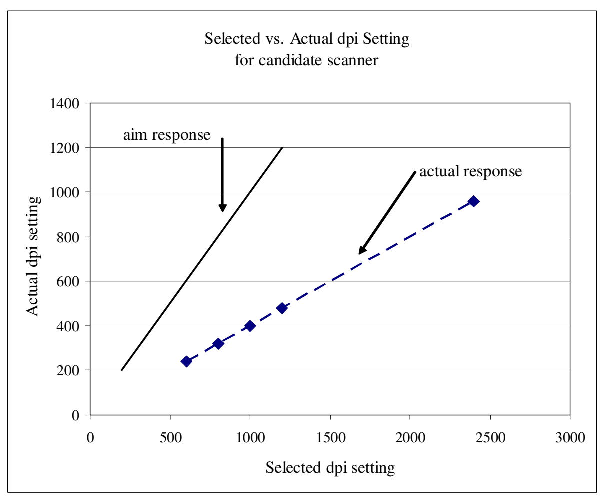

***Figure 6**: Actual vs. selected dpi error for a selected scanner*

Generally, two statistics are monitored for a series of measurements on a given variable (e.g. resolution, tone, magnification). They are the mean level and variance of a process,

and are associated with accuracy and precision respectively. For digital image capture, either one or both of these may be important. For instance, for tone and color control, the precision, or repeatability, of a process is likely much more important than the accuracy of hitting a selected aim point. This is because mean level tone/color inaccuracies in digital images can usually be processed back to a given aim as long as the error is relatively constant and a
record of the aim is available (i.e., the *Golden Thread* target). If, however, the tone/color capture is, on the average accurate but suffers from high variability, the downstream image quality will be chaotic and probably generate complaints. Monitoring, preventing, and preparing for such possibilities are the essence of SPC.

Database Population

> *I find that a great part of the information I have was acquired by looking up something and finding something else on the way.*
>
> *~Franklin P. Adams*

Unusual sports statistics, purchasing recommendations from your favorite online retailer, or even auto recall notification are all enabled by a database. Generally these databases have served to make many of life's daily tasks easier, or certainly more informed. Similarly, SPC is able to realize the benefits of databases as a component of a digitization workflow. Quality procedures that make use of image targets, image quality review, file format validation tools, all generate data. And while that data can be used in real-time to ascertain the quality of a singe object, the real benefit to using the data is in improving the digitization process as a whole. This is value that SPC brings to an organization in terms of reduced cost and improved quality. Rather than detecting a bad image at the end of the process, requiring re-scanning, the goal is to obtain data throughout the workflow to detect problems early, and even prevent the problems from reaching the point of failed products.

To do this, the data must be recorded from many points and centrally stored in a database system. The database allows the data to be represented in many different ways for a variety of uses. In SPC, the primary representation is a control chart, and those charts are used for benchmarking, process monitoring, trending and prediction. The feedback obtained from the process monitoring can also be used to establish procedures that dictate how often monitoring needs to be performed.

In a typical digitization workflow, a number of steps in the process can be monitored and the data written to a database. Let's imagine a model workflow originating at a vendor, with tools used to generate the following types of data for SPC:

- MD5 Checksums verifies that the product created is the product received
- JOVE verifies that the files are valid, well formed, and to specification
- A device-level target verifies that vendor scanning devices are generally performing to the level of quality required by contract
- An image-level Target verifies that individual images conform to specification

From this data, we can not only catch errors at any point in the process, but develop models around trends. Trend analysis can be used to prevent problems and improve processes. Here are just two examples of using trend analysis in a digital workflow:

A) Resolution data as a measure of spatial frequency response (SFR) from a device-level target continues to fall within an acceptable quality level (AQL), but the mean daily values are dropping. What's more, they are drifting more from the corners than from the center target. Before the levels fall below the AQL, the problem is diagnosed and corrected at the scanner. In this case, a problem with the book cradle was slowly changing the angle of the cradle and it was no longer parallel to the plane of the sensor array.

B) The measure of SFR from a image-level target exceeding the AQL on a lengthy project involving rare books. The measures are inconsistent and no pattern is readily evident. The problem is traced to a particular operator who had a higher frequency of errors as a result of not following focusing procedures for the scanning equipment being used.

In each case, the data provided information to correct the cause of the problem causing bad images rather than simply correcting the image by re-scanning. We are not suggesting that problem analysis does not take place in current scanning operations. There is a high level of dedication to quality that exists at institutions involved in scanning cultural heritage materials. The methods described here simply provide tools and procedures to achieve the highest level of quality for a given level of effort.

Conclusion

> *Laws control the lesser man.*
> *Right conduct controls the greater one.*
>
> *~Chinese Proverb*

Practitioners of digital imaging in the cultural heritage community are becoming increasingly savvy. They are realizing that digital imaging, just like traditional film imaging, suffers from its own set of unique image quality problems. The difference is that for digital imaging, the onus of quality control has largely trickled down to the user. The unprecedented freedom and variety of choice we enjoy comes with a price of responsibility, especially as it relates to quality.

Most imaging centers are not averse to assuming this responsibility, but are lacking the tools to enable the execution of robust SPC practices worthy of ISO certification. As demonstrated here, targets and analysis software are the keys to this enablement. No invention is required, just integration, hard work, and the enthusiasm to provide quality images that will endure as true archives.

Acknowledgements

The authors would like to thank Dr. Peter Burns (Eastman Kodak) and Ron Murray (Library of Congress) for the helpful suggestions and material for this paper. Special thanks also to Phil Michel (Library of Congress), Rhea Garen (Cornell) and the staff at Stanford's Green Library for supplying information rich sample images for image performance analysis in the context of real workflows.

References

[1] Walter A. Shewhart, Economic Control of Quality of Manufactured Product, Amer Society for Quality (1980), ISBN 0873890760.

[2] http://www.i3a.org/iso.html

[3] R. Murray, Managing a Quality Digitization Practice in Cultural Heritage Institutions: Statistical Quality Control Tools and Techniques, Proc. IS&T Archiving Conference, (2006).

[4] Freddy Balle, Michael Balle, The Gold Mine: A Novel of Lean Turnaround, Lean Enterprises Inst Inc (2005).

Author Biography

Michael Stelmach has twenty-five years of information management experience, with a concentration on providing digital access to content originating from print. Formerly the Vice President of eBook Production at netLibrary, he is currently the Manager of Digital Conversion Services in the Office of Strategic Initiatives at the Library of Congress

Don Williams is a frequent technical contributor/instructor on digital image capture metrology, image quality, and image fidelity issues in the scientific and cultural heritage communities. He is an image scientist at Eastman Kodak Research Laboratories in Rochester, N.Y. and is a participating member and technical leader for ISO / TC42 digital imaging metrology standards. He is currently the editor for ISO 12233, Digital Camera Resolution standard, and co-leader for ISO 16067-1, Digital Scanner Resolution standard.