ARCHIVING CONFERENCE
IS&T imaging.org

https://doi.org/10.2352/issn.2168-3204.2017.1.0.165

A Decade of Experience with Digital Imaging Performance Guidelines: The Good, the Bad, and the Missing

Don Williams, Image Science Associates, and Peter D. Burns, Burns Digital Imaging

Abstract

As with any initiative, despite design intentions, the first efforts have unexpected positives and how-did-we-miss-that *negatives. The FADGI (Federal Agency Digital Guideline Initiative) and Metamorfoze guidelines are no exceptions. Whenever such efforts are brought to practice in the field we learn. And that is good, because it forces behavior, software, and hardware to evolve to be more resilient. We address these developments in digitization for cultural heritage collections. We look-back on the last decade of our experience with these guidelines and discuss progress, limitations, and future directions.*

Introduction

In general, the FADGI[1] and Metamorfoze[2] digital imaging performance guidelines serve their purpose well for manufacturers, vendors, and collection custodians. However, they are, after all, guidelines. They provide guidance, not necessarily, absolute imaging specifications. They are intended to identify levels of reasonable agreement for good digital image capture, particularly for the high-volume workflows in the cultural heritage sector. In many ways, they have succeeded.

Based on our experience in the implementation of these imaging guidelines for various institutions, we share our observations with the reader; good, bad, and in-between. The discussion will be drawn from actual examples of field practice. Our intent is to improve future versions of the guidelines, and the imaging practices meant to comply with them.

We start with the history of the guideline documents themselves as they moved from somewhat inconsistent and confusing set of rules to their present image-science based architecture. We emphasize that many of the failures are not so much the fault of the users, but result from the clashing of theory and practice when implemented.

Initially, an overlooked challenge to the broad implementation of the guidelines was the assumed digital imaging knowledge of the people meant to use and comply with them. This image-literacy gap needs to be narrowed for further success of the guidelines. Solutions for closing these gaps and other will be offered.

A Little History

Metamorfoze and FADGI were born at approximately the same time a decade ago. The motivation for these initiatives was to help manage the range of digital image capture practices across cultural heritage institutes, and to establish nominal levels of image-capture performance. Simply agreeing to a set of guidelines by a committee was insufficient though. The guidelines had to also be scientifically sound, and based on vetted international standards. If this was to be done, it needed to be done correctly. A common terminology was the first goal.

It ain't what they call you; it's what you answer to

— W. C. Fields, American comedian, actor and writer

Image-capture terms such as resolution, sampling rate, gamma, white balance, and color fidelity have always been bantered about in describing imaging performance. However, such imaging-jargon terms need definition and explanation to be useful to the wider community. Borrowing from imaging performance standards of ISO/TC42/WG18 and IEC, several of these imaging characteristics can be confidently measured, and with little ambiguity, used indicate imaging goodness.

It is important to mention that these guidelines apply strictly to image capture and not necessarily to the rendering or display of the captured data. However, a virtual-display environment is assumed for output referred color-spaces. The philosophy is to acquire robust, digital images that are reusable for a variety of applications, i.e. without the need to re-digitize. A resilient digital image object can be rendered for display for multiple applications such as publications, research, access, or other artistic intent. The guidelines enable this.

Finally, let's clear the air. There continues to be a perception that the FADGI and Metamorfoze guidelines are somehow opposed and are a mono-culture in themselves. Not true. Both initiatives use the same set of ISO standards for measuring imaging performance, and agree on aim and tolerance levels for 95% of their adopted metrics. For our purposes in this paper these are applied equally in both. Although there are some differences they largely apply to their flexibility in field use. In the future these differences may be greater. As of this writing no further developments are underway for the Metamorfoze guidelines. FADGI does continue to add changes to their documents, and expects their guidelines to improve with future user feedback and experiences.

Good Targets Make Good Tools

The catalyst for any of the guidelines is a ruler (reference object) by which imaging performance can be measured. Resilient, comprehensive, and well-characterized reference targets enable this. While targets have often been used for past digitization efforts they have sometimes been treated more as talismans, rather than tools for quality control, remediation, and benchmarking. Users might feel that by placing a ragged uncharacterized printed target in the field of view, they are secure in exercising good imaging practices. They often miss the key element of actually using them for resolution, exposure and color assessment. One needs to exploit

the targets for them to be of any benefit. Today this means analysis software.

In Fig. 1 we show an outdated test target being used in a book during book scanning. The target's test patches were not (color-) characterized. In addition, the target was not in the same plane as the page being scanned. It was about 2 cm. lower. The visible margin between target and book is not a support, but rather the gold-leaf edges of 100 or so pages. The scene is, *The Approach to Karnac* from Ref. 3.

Figure1: Example of poor usage of an outdated imaging test target

Along with commercial and free easy to use software to do so, true analytical analysis is now possible and frequently practiced by using the targets for insight. Some highlights of these analyses are provided below.

Resolution and sampling efficiency: It is quite easy to distinguish between resolution quantity and resolution quality by evaluating the sampling efficiency[4] of a capture system. This normalized resolution metric has been especially helpful with film scanning where high sampling rates (greater than 2000 dpi) are frequently called for but cannot be easily met with most scan devices.

Exposure, Gamma, White Balance, and Gain Modulation: These metrics are derived from the gray scale target features. Since these are the foundation of any standardized color space compliance it is important that they be analytically verified, especially across the full scale of neutral tones. It is important to understand that these target features need to be spectrally neutral, and not just visibly neutral as they are in photographic color targets. This has been an important lesson, and why we often discourage users from using standard color photographic targets for creating color profiles.

Standardized Color Space Compliance: The cited guidelines have encouraged users to test compliance with respect to standard color spaces. By using established color-difference (ΔE) formulas, judgments of both individual and average color errors are computed and comparable between systems. In addition, we have learned that maximum ΔE values greater than 6.0 have little relevance since such values were not in the scope of the original work used to derive the ΔE formulas.

Setting different ΔE values as part of the guidelines needs to be reconsidered. Perhaps a substitute for the maximum ΔE metric should be a ΔE-dispersion metric. For instance, rather than testing for ΔE_{max} a more suitable measure would be a 90% percentile measure. This approach is already reported by the BasICColor Input color profiling software, and could be a model for a substitute to ΔE_{max} called for in the guidelines.

The test targets advocated in FADGI and Metamorfoze guidelines are well designed for broad imaging performance assessment and have been perhaps the most important component in lowering the barrier to imaging performance assessment. This is a certainly a significant benefit of the practices developed by these two initiatives.

As of this writing, only the FADGI guidelines address film or transparency film digitization. These are outlined for both microfilm and standard film formats (e.g., 35 mm, 60 mm, and 4 in. x 5 in). High-resolution grayscale targets for these are available, but are not yet defined for color transparencies.

Two cautionary notes on the subject of targets need to be considered. These were identified after a number of field reports from users were reported. These are discussed below and are some of the areas that need improvement.

Context matters. The color patches used in the current target sets are largely based on legacy colors derived from consumer imaging applications. These color sets are not always well suited for cultural heritage imaging applications, specifically subtle pastel colors. So, while it may be possible to obtain good color fidelity performance on the target patches of the existing targets, the actual colors in the collection content may suffer. In the future color fidelity assessment, should separate calibration targets from validation targets to better assess color imaging capture performance.[5]

Another area for improvement is to adopt target-specific reference files when evaluating imaging performance. These are aimed at both the neutral and color target patches. We have found that assumptions on the default target colorimetry can be significantly different from the actual measured values, especially for the high and low L^* samples. An example of the annoyance that this introduces is told below.

> A large institution required its contractors submit an image of a test target daily for image quality verification. Their criterion for acceptance was a specified L^* difference between two gray patches on a popular color-calibration target. The institution based this criterion on batch (nominal) L^* values for those two patches and not on the specific values of the actual test charts being used. This required the contractor to unnecessarily sort targets whose results fell within the tolerance, in order for image batches to be accepted. All of the scans in this case were acceptable but were rejected because the target values were not accommodated. The proper approach would have been to build acceptance criterion based on specific target data rather than batch data.
>
> *A good man always knows his limitations*
> — Clint Eastward as Harry Callahan, *Magnum Force* (1973)

The lesson: there is going to be random variability at all stages of the digitization process and some strategies for dealing

with it need to be considered, so that good product is not rejected because of an intolerance to hard acceptance thresholds. This is one area that needs to be addressed with the current guidelines.

On a related note, too often the measured performance levels fall outside of the published guideline values by very small amounts (e.g., 1-2%) which could easily be attributed to measurement error. We suggest use of a *caution zone* for such cases. As indicated earlier, these are guidelines not absolute laws. They are intended to be flexible and reasonable recommendations. Perhaps one philosophy to remember here is that *precision often trumps accuracy*. As long as there is a reasonable level of accuracy to any of guidelines tiers, the more important goal is to ensure that performance consistency is achieved.

Image Literacy and a Community of Use

Both the Metamorfoze and FADGI guidelines for still imaging are technical by nature. While this is an admirable and righteous objective, most members of the community who are meant to practice these guidelines are not of that ilk. Even trained photographers are challenged by some of the technical aspects of the guidelines. In the enthusiasm of creating the guidelines perhaps a look back at those who were to practice them was neglected. So, one area that needs improvement is the technical training for the community (vendors, manufacturers, and end-users) meant to follow these guidelines.

While conference workshops and lectures by experts on the subject are (ahem) helpful, additional resources are needed. As examples, efforts at Lyrasis[6, 7] and Univ. of Michigan's have helped in this regard. Lyrasis has posted five theory-to-practice online tutorials[8] Paul Conway at the University of Michigan School of Information[7] has included coursework on the FADGI guidelines in his teaching curriculum. The work under the ISO/JWG26 is another emerging resource.

Color, Colour, Kolor

One example of where user training would be helpful is in awareness of colorimetry. Currently, a source of confusion in the guidelines is in the reporting of color and white balance performance metrics. Technically, these metrics should be reported as colorimetric L^*, a^*, and b^* components. This is the approach in Metamorfoze. The one practical problem with this approach is that very few practitioners are comfortable or even familiar with colorimetric terminology. How does one translate between the RGB- centric approach in camera and scanner user interfaces into required $L^*a^*b^*$ values? This is problematic when trying to remediate non-compliance issues.

FADGI began cautiously on this subject by defining guidelines in terms of image RGB values. Admittedly, this can be ambiguous because the meaning depends on the color-space chosen, but was considered acceptable for the initial version. The thinking was to begin by favoring simple recommendations, then transition toward more technically accurate colorimetric methods as needed. Future versions of the FADGI guidelines are likely to adopt colorimetric aims and tolerances for the four-tiered performance levels.

Alternative Arte-Facts

Another area for improving digital image literacy is in recognizing digital image artifacts. These are many and varied and are not easily measured analytically. They defy objective measurement and can usually only be detected reliably by visual inspection of selected target features. Two examples of these are shown in Figs. 2 and 3 for the hyperbolic resolution wedge used in many of ISO resolution standards. Many of these are technology specific. Just knowing what type of scanner technology is used allows one to anticipate certain artifacts.

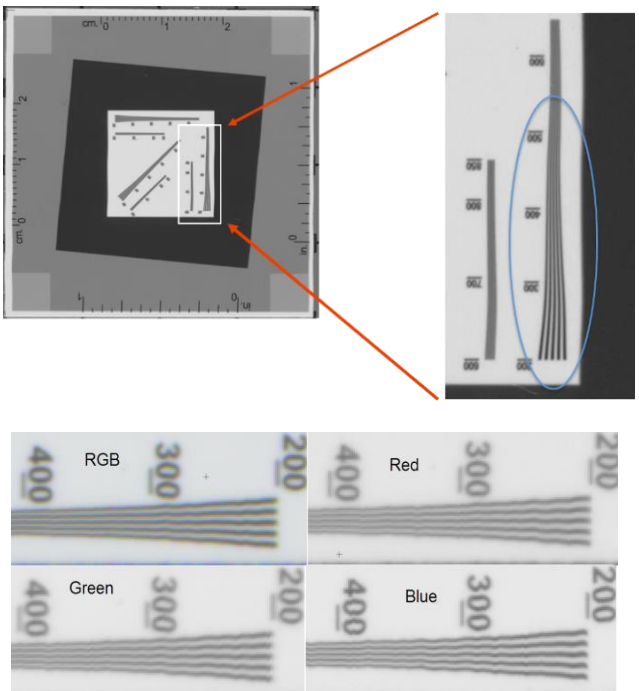

Figure 2: Examples of wavy line artifacts (circled) due to scanner velocity fluctuations. Often separating the image into its color components makes the artefacts more obvious.

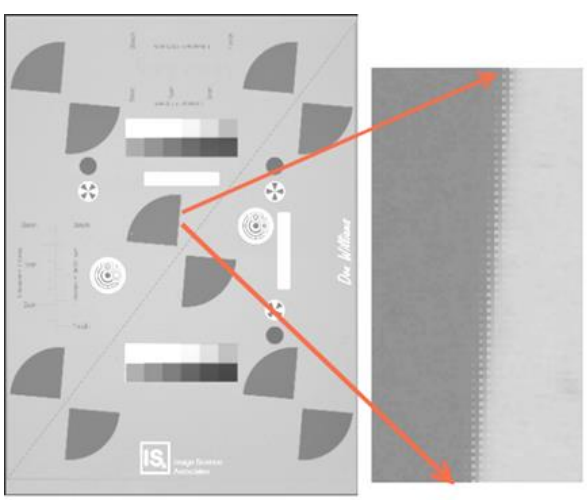

Figure 3: Checkerboard artifact along an edge from a 4-shot digital camera with poor data reconstruction.

Consistency in Image Manufacturing

To date, efforts on guideline compliance have focused largely on demonstrating single instances of selected imaging performance levels. Given the sheer size of the collections digitized by cultural heritage institutes it is no exaggeration to view the associated workflows in the context of a manufacturing process where consistent good imaging is required for large volumes of images, not just one. The greater goal then is to maintain these levels over the course of a large project, often with tens or hundreds of thousands of images captures over several months. Managing the inescapable variability that will occur for such large projects needs to be embraced. Sustaining consistency through quality control is very important and makes project management and image reuse so much easier.

We introduce this idea with an anecdote from *Understanding Variation: The Key to Managing Chaos*.[9]

Some Days are Better than Others [9]

Statistician, David Chambers, found a graph (Fig. 4) on the office wall of the president of a shoe company. The vertical-axis label was 'Daily percentage of defective pairs.' Intrigued, David asked the president why he had this graph on the wall. The president condescendingly replied that he had it there so he could tell how the manufacturing plant was doing. David responded with, 'Tell me how you're doing.' He paused, looked at the chart on the wall, and then said,

'Well, some days are better than others!'

Although he had the data, he had no way of analyzing the values and interpreting results.

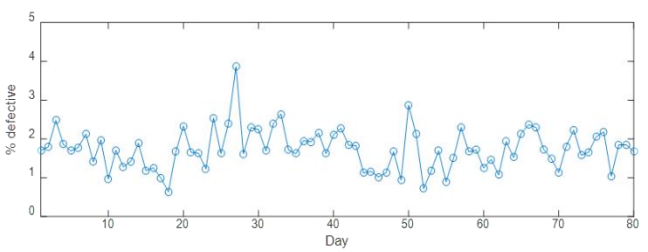

Figure 4: Typical Defect vs. daily control chart

Several institutions already adopt practices that enable such control charts by requiring submission of periodic, daily or bi-daily, image quality target scans to audit imaging performance levels.[10,11] With hot folders and scripted analysis software such quality control practices can be easily implemented. Future versions of the FADGI guidelines are likely to include suggested strategies on the level of detail of this workflow monitoring.

Delay not, Caesar. Read it instantly

— Artemidorius in *Julius Caesar* 3, 1 W. Shakespeare

Though large amounts of image performance data are available to analyze, only key, high priority, metrics need to be monitored at first; the ones most likely to vary. Exposure is one of these. In Fig 5 an example control chart derived from alternating-day submissions of a target from a robotic book scanner is shown. Variations in apparent exposure can likely occur from day-to-day, and for recto/verso views.

Fig. 5 illustrates how only three gray levels for each page view are tracked to reveal that a significant change in exposure has occurred about halfway through the project. The three gray level values for recto/verso captures track well however. They are so close that they lie virtually on top of one another (e.g., blue/recto and black/verso plots from the white-patch values).

Since all the gray levels diminished in count value from the initial history an under exposure occurred. This is revealed with the images of Fig. 6 that are linked to the anomalous control chart behavior. The image on the right is notably darker.

After analyzing the test dates, under exposures occurred after a long holiday weekend. Variations like this are common whenever workflow transitions occur, e.g. after a long shut down. Changes often occur when personnel, equipment, or shift-work changes are made. Experiences like this show just how beneficial the guidelines can be in helping to identify unacceptable imaging performance changes, to provide a consistent product.

It is worthwhile to note that either the left or right image of Fig. 6 may be acceptable as singular examples. However, in the context of the greater project the under exposed results indicate an inconsistent process, project management more difficult. Ideally, the type differences demonstrated in Fig. 6 should have been detected before any further captures were made, thus avoiding any rework and cost later.

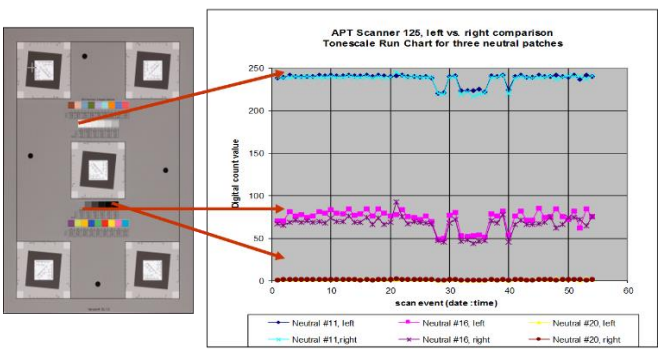

Figure 5: Bi-daily control chart for a robotic book scanner

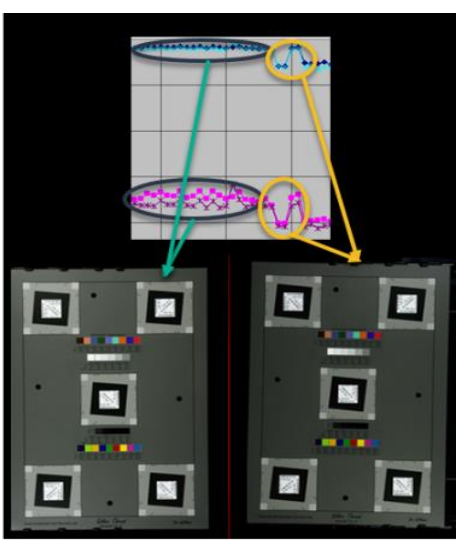

Figure 6: Expanded view of control chart showing exposure differences

Adopting from Standards

In the future, we will likely see the imaging performance guidelines incorporate missing imaging characteristics. Consistent with their previous development, we would expect the adoption of methods based on international standards.

Optical Distortion

As an example, consider image distortion, such as that normally due to a camera lens. Common examples of this are pincushion and barrel distortion. There are two recently released ISO standards that can be applied. ISO 17850[12] provides a geometric distortion measurement, and ISO 19084[13] addresses the wavelength (color) dependent nature of optical distortion. The evaluation of both types of image distortion can be accomplished using a test target with an array of equally-spaced dark circles.

Further description of these evaluation methods is given in Ref. 14. Results for the two analyses can be reported in terms of % distortion and color displacement values. In addition, the computed distortion can be presented as an image-wide field, to facilitate diagnosis of failing performance. An example is shown in Fig. 7.

Conclusions

The FADGI and Metamorfoze digital imaging performance guidelines serve their purpose well. Using methods largely based on international imaging standards, they have been used to qualify imaging service providers, equipment acceptance and set-up, and process control. Though there are limitations, we have described the ways in which these will likely be mitigated in the future. These include,

- continued use and development of multi-test test targets as reference objects
- development of color targets for film scanning
- addition of several imaging performance measures, such as distortion and chromatic displacement

In addition, we can expect adaptation and adoption of guidelines to 3D imaging applications. For example, 3D- projection image capture.[15]

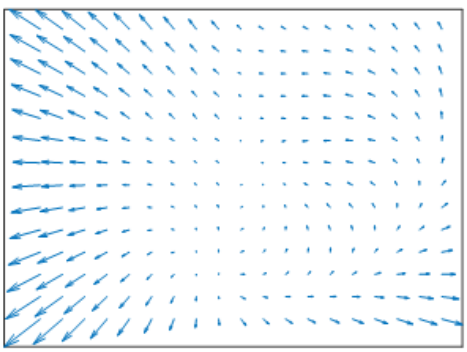

Figure 7: Quiver plot representing measured geometric distortion where each arrow (distance) length is drawn as 150%, from Ref. 14.

Acknowledgements

The authors thank a host of contributors over the years for their suggestions and shared experiences. These include Tom Rieger and staff at the Library of Congress, Hans van Dormolen and other staff members at the KB in The Hague, Dietmar Wueller, Scott Geffert, digitization staff at Smithsonian Institute, Picturae, and Digital Transitions.

References

[1] FADGI Still Image Working Group, ed. T Rieger,, US Library of Congress, 2016 http://www.digitizationguidelines.gov/

[2] H. van Dormolen, Metamorfose Preservation Imaging Guidelines, Image Quality version 1.0, Nat. Library of the Netherlands, 2012. https://www.metamorfoze.nl/sites/metamorfoze.nl/files/publicatie_documenten/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

[3] L. W. Yaggy and Y. L. Haines, The Approach to Karnac, Museum of Antiquity, Standard Pub. House, Chicago, pg. 385a, 1881.

[4] P. D. Burns and D. Williams, Sampling Efficiency in Digital Camera Performance Standards, Proc SPIE 6808, 680805, 2008.

[5] D. Williams and P. D. Burns, Rethinking Image Color Correction, Validation and Testing, Proc. Archiving Conf., IS&T, pg. 175-180, 2016.

[6] Lyrasis, https://www.lyrasis.org/Pages/Main.aspx

[7] Lyrasis, Intro to FADGI for Still Image Digitization, on-line tutorial: https://lyrasis.adobeconnect.com/_a782640538/p89tcok4mwc

[8] P. Conway, Enhanced Education for Better Imaging Practices, Univ. Michigan website, http://www.academia.edu/2796022/Enhanced_Education_for_Better_Imaging_Practices_A_Case_Study_at_the_University_of_Michigan

[9] D. J. Wheeler, Understanding Variation: The Key to Managing Chaos, SPC Press Inc, pg.15-16, 1993.

[10] M. Stelmach and D. Williams, When Good Scanning Goes Bad: A Case for Enabling Statistical Process Control in Image Digitizing Workflows, Proc. Archiving Conf., IS&T, pg. 237-243, 2006

[11] R. J. Murray, Managing a Quality Digitization Practice in Cultural Heritage Institutions: Statistical Quality Control Tools and Techniques, Proc. IS&T Archiving Conference, pg. 96-104, 2006.

[12] ISO 17850:2015, Photography – Digital cameras – Geometric distortion measurements, ISO, 2015.

[13] ISO 19084:2015, Photography – Digital cameras –Chromatic displacement measurements, ISO, 2015.

[14] P. D. Burns and D. Williams, Going Mobile: Evaluating Smartphone Capture for Collections, Proc. Archiving Conf. IS&T, pg. 102-106, 2016.

[15] P. D. Burns and D. Williams, Evaluation of 3D-Projection Image Capture, Proc. Archiving Conf., IS&T, pg. 70-73, 2015.

Author Biography

Don Williams is founder of Image Science Associates, a digital imaging consulting and software group. Their work focuses on quantitative performance metrics for digital capture imaging devices, and imaging fidelity issues for the cultural heritage community. He has taught short courses for many years, contributes to several imaging standards activities, and is a member of the Advisory Board for the interagency US Federal Agencies Digitization Guidelines Initiative, FADGI.

Peter Burns is a consultant supporting digital imaging system and service development, and related intellectual property efforts. Previously he worked for Carestream Health, Eastman Kodak and Xerox Corp. He is a frequent conference speaker, and teaches courses on these subjects