Image Quality Assurance for the Real World

Paul W. Jones and Chris W. Honsinger; Certifi Media Inc.; Rochester, New York, USA

Abstract

The need for quality assurance in image digitization programs has long been recognized by the cultural heritage and digital archiving communities. To that end, detailed quality guidelines, such as those from NARA and Metamorfoze, have been produced and shared within these communities as an essential first step in driving the adoption of i) quantitative quality metrics and ii) common approaches to the calculation of such metrics. These quality metrics are fidelity-based in that they address the question: Is the digitized image an accurate representation of the original content or object? Examples of such fidelity-based metrics are spatial frequency response (SFR) to assess sharpness, opto-electronic conversion function (OECF) to assess tonescale reproduction, and flat-field standard deviation to assess noise. These measurements are produced from captured images of one or more test targets, which serve as reference input signals.

While these efforts are important and necessary, they are not entirely sufficient when it comes to implementing quality assurance programs in real-world production environments. In this paper, we'll discuss the next steps that are required to make quality assurance more than an academic exercise. These steps include fully automated, real-time quality analysis that integrates seamlessly and easily into a digitization workflow; simple and convenient quality metadata management; efficient exception handling to identify and fix quality problems; and quality metrics that go beyond current fidelity-based metrics. Our company, Certifi Media, develops technology and solutions to specifically meet these real-world quality assurance needs. In this paper, we'll discuss how a practical quality assurance program can be implemented with such technology and solutions.

Historical Perspective

For well over a decade, the cultural heritage and archiving communities have recognized the need for improving image digitization practices so that valuable records can be adequately represented in digital form. To this end, various government agencies, academic institutions, and other organizations and consortiums have published an array of image digitization guidelines and best practice recommendations [1].

Among these publications, the NARA guidelines [2] and Metamorfoze guidelines [3,4] stand out as providing clear information on quantitative quality metrics and related tools that can be used in setting up a digitization workflow with a desired quality level and subsequently monitoring the quality performance over time. Both of these guidelines leverage standardized methods for objectively measuring key image quality attributes, such as sharpness, tonescale reproduction, color reproduction, and noise. Examples include the ISO standards for measuring spatial frequency response (SFR) using slanted edges, the opto-electronic conversion (OECF) using neutral density patches, and noise using uniform, flat-field patches.

The quality metrics and measurement methods in the NARA and Metamorfoze guidelines are fidelity-based, that is, they are aimed at answering the question: Is the digitized image an accurate representation of the original content or object? The assessment of fidelity is facilitated through the use of one or more test targets, which serve as reference input signals that can be compared to degraded output signals. This is a very powerful approach to image quality assessment, but it is not necessarily the only approach as we'll discuss later.

The adoption and promotion of quantitative quality metrics is extremely important in making robust and meaningful quality assurance available throughout the heritage and archiving communities. A very real benefit of these guidelines and best practices is that they provide a common language for people to discuss and understand the need for quality assessment and how it might be put into practice.

However, this activity represents only the first step in driving the widespread usage of quality assurance programs in everyday digitization operations. The goal of this paper is to outline the critical next steps and discuss how they can be accomplished.

Enabling Practical Quality Assurance

A practical quality assurance (QA) program must meet three basic requirements:

- Provide meaningful information;
- Be convenient to use; and
- Be cost-effective to implement.

A QA program that doesn't meet all of these requirements will only be partially implemented or will not be implemented at all. A common fallback position is to perform the QA by having people review images, which certainly is not cost-effective.

The guidelines and recommendations from NARA and Metamorfoze only address the first point, providing meaningful information through specific quality metrics and measurement techniques, specifically for the case of fidelity-based QA. However, even with these guidelines, there is further clarity that is needed with regards to what is and isn't meaningful in the context of a QA program. We'll first discuss this issue, followed by the steps needed to make QA convenient and cost-effective.

Making QA Meaningful

Capability assessment vs. performance monitoring

It is worthwhile to make a distinction between capability assessment and performance monitoring [5]. Capability assessment is about determining the peak quality that an imaging system can produce, while performance monitoring is about the distribution of quality that is achieved in day-to-day practice. A simple example of this distinction is illustrated in Figure 1.

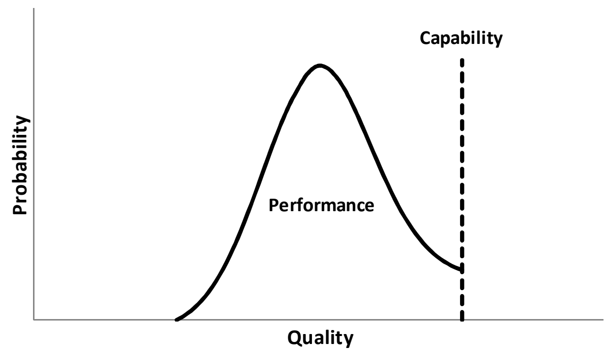

Figure 1. Example of capability (peak quality) vs. performance (quality probability distribution).

Capability assessment is essential when equipment for a digitization program is being purchased, set up, and calibrated. One needs to know that the overall system is capable of meeting all digitization program needs, such as the maximum dpi (both real and effective), the noise level at different input densities and ISO settings, the overall dynamic range of the capture device, the color reproduction over a large gamut of color patches, the peak misregistration of the color channels, geometric distortion, etc. A rigorous and full assessment of capability may be time-consuming and involve a variety of measurement tools and techniques, but that is acceptable because it is a process that is done infrequently.

On the other hand, performance monitoring is something that is practiced on an on-going (and preferably very frequent) basis. The focus is not on peak quality, but rather on the distribution of quality that is achieved in real-world production. As such, performance monitoring involves tracking those quality attributes that are likely to change, whether due to equipment variability, operator error, or another cause. A quality assurance program is really about performance monitoring, with capability assessment being a necessary precursor to an on-going QA program.

Focus on what matters

The point of these comments regarding capability versus performance is that there is often a blurring of the two in quality guidelines, which makes it difficult for the non-expert to distinguish what quality measurements are essential as part of an on-going quality assurance program.

For example, there is little or no benefit in continuously assessing something like pin cushion or barrel geometric distortion as that component of quality is not likely to change. However, there may be significant benefit in assessing whether the capture sensor and book or manuscript are co-planar in a copy stand setup as misalignment (which induces a perspective distortion) can occur fairly easily during routine operations.

Likewise, assessing color reproduction using dozens of color patches during QA is more confusing than helpful because typical color reproduction problems (such as poor white balance or an incorrect ICC profile) can be detected with only small number of color patches. However, a full color analysis should be part of the capability assessment process, where one or more ICC profiles are calculated for use in the actual production environment.

Set realistic quality goals

One of the most difficult aspects of a meaningful quality assurance program is setting realistic quality goals. Some digitization guidelines specify quality levels that are very high, and it is crucial to make a conscious decision that the benefit is worth the cost for a given application. While the various digitization guidelines and recommendations can provide useful starting points, it is more likely that setting the quality goals for a specific application will require fine tuning and perhaps even ignoring certain quality recommendations if warranted. Blind acceptance of any quality guidelines is not in the best interests of either content owners or digitization service providers. Rather, careful consideration must be give to intended purposes of the digital images and the required level of quality. For this reason, education and experience will remain essential elements in driving the adoption of meaningful quality assurance programs.

Making QA Convenient and Cost-Effective

While specific quality metrics, measurement techniques, and guidelines may change over time (and even from job to job), QA will not become part of the day-to-day digitization operations unless it is convenient and cost-effective. At a minimum, achieving this goal requires the following:

- QA is integrated seamlessly into the digitization workflow, rather than being treated as a separate process;
- The quality metadata is packaged for convenient access and review at any time; and
- Improved exception handling is available to manage and correct quality problems.

Integrate QA seamlessly into the workflow

Quality assurance is often treated as an afterthought and viewed as an inconvenience. To overcome this current reality, QA must become an integral part of the digitization workflow, with little burden to the operator and minimal processing overhead.

Automated target detection and analysis

At the core of QA integration is full automation of the test target detection and analysis. Fortunately, this is straightforward to accomplish with proper test target designs and image processing tools. As examples, two Certifi Media reflection test targets (FC-1 and FC-2) are shown in Figure 2.

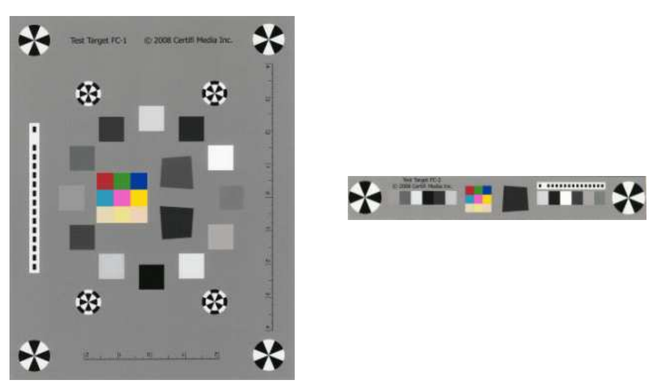

Figure 2. Certifi Media test targets with fully automated detection and analysis (left) FC-1 target, 8.5" x 11"; (right) FC-2 target, 1.25 x 11".

Using a combination of machine-vision target elements and flexible XML target descriptors, these targets can be automatically detected and analyzed with absolutely no operator intervention. The analysis includes a full suite of quality metrics, including sampling rate (DPI), sharpness (SFR), tonescale (OECF), noise (std. dev.) color and neutral reproduction (delta E), and uniformity. Each target also includes a simple bar code that specifies the manufacturing lot number, allowing each lot to be calibrated for manufacturing variations in the aim densities and colors, which can be taken into account automatically during analysis.

Using these targets is very simple in an actual workflow. The robust detection and analysis capability allows them to be used in any orientation and location within the image frame. The larger target is designed as a device-level target to be captured intermittently, while the smaller target is designed as an image-level target to be used alongside content and captured with each image [6]. The targets are also designed to capture the same information so both can be used within a single job without confusion.

The need for speed

While automation is the first step in QA workflow integration, it must also be combined with rapid processing. The benefits of automation are largely negated if it takes minutes to detect and analyze a target. Through careful algorithmic designs and low-level image processing optimization, we're able to do a full detection and analysis cycle of the Certifi test targets in approximately 300 msec using a common dual-core computer, with little dependence on the image size. Faster speeds can be obtained with quad-core processor owing to a multithreaded design.

Even with fast, automated target detection and analysis, it is still necessary to simplify the QA process as much as possible by using only a single target for all quality measurements. We have seen RFPs for digitization projects that specify as many as five different targets as part of the QA process, which would still be burdensome for an operator even with automated detection. However, because a single target has limited real estate, it is necessary to restrict the measurements to the essentials that are needed for proper QA (refer to the previous section entitled 'Focus on what matters").

The power of having simple, fast, automated detection and analysis is that quality monitoring can now be performed in near real-time and on a continuous basis. Test targets can be captured and analyzed efficiently and conveniently at any time, not just a few times per day. This allows quality problems to be detected and corrected much sooner, minimizing the amount of costly rework.

Quality assurance and image processing go together

An image digitization workflow invariably involves image processing beyond the original capture process. Because image processing often affects one or more attributes of image quality (sharpness, tonescale, noise, color reproduction, etc.), it is a natural fit to combine QA as part of the image processing. Moreover, digitization programs often spend considerable time and effort in optimizing the overall workflow, and a QA tool that requires a separate software application outside of the image processing workflow is inefficient and inconvenient.

An example of an application that integrates QA and image processing is shown in Figure 3. This interactive tool ("Script Builder") is part of the Certifi Pedigree QP product, which also includes a batch mode for integrated image processing/QA.

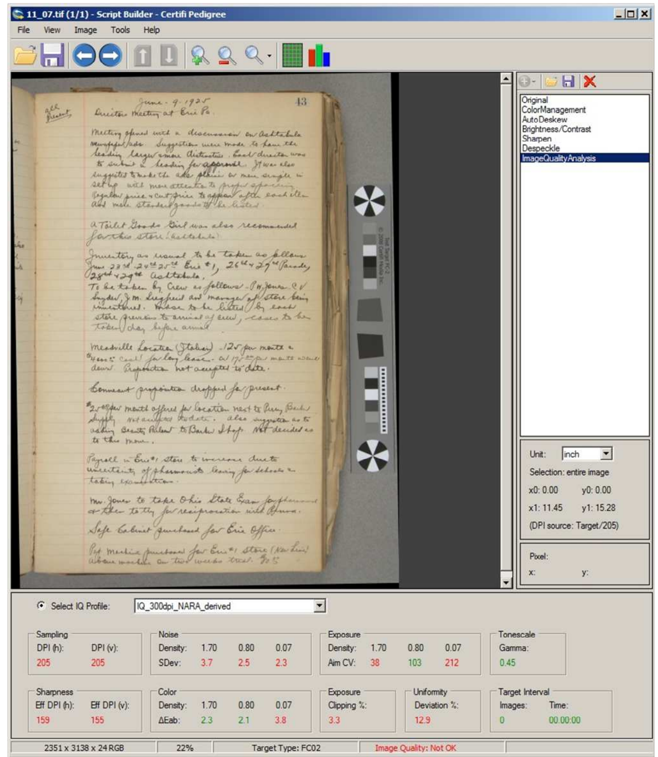

Figure 3. Certifi Pedigree Script Builder – an Interactive tool that combines image processing and QA.

As processing steps are added, removed, or modified with this tool, the image quality analysis is automatically updated in the display panel at the bottom. The quality metrics are compared against a desired set of specifications (a "quality profile"), and the quality metrics are color coded red or green to denote whether they are in or out of spec. In this way, an operator can quickly design an image processing script that meets the quality specifications for a single image or for a set of images with similar characteristics.

Another benefit of integrating the image processing and QA is that having access to a test target allows for automated image processing algorithms that can use the well-characterized physics of the target to set the tonescale, sharpness, noise, etc. to desired set points directly [7].

Finally. integration of QA into existing workflows may also require repackaging of target detection and analysis technology into SDK or DLL solutions so sophisticated users can access the core methods from other software applications or scripting languages. This is a current development focus for Certifi Media.

Manage the quality metadata

Quality metadata packaging and access

The measurement and display of quality metrics during image acquisition and processing is obviously useful in monitoring and improving the production process, but what about images that are archived? Even at a future time, it may be highly useful to have the quality metrics available for review. For example, images may be acquired originally for a certain application and then repurposed at a later time for another application with different quality requirements. If the quality metrics are readily available, it's a simple matter to compare them against a new quality profile to determine if the quality is sufficient for the new application.

The most convenient and error-free way to package the quality metrics for robust access and retrieval is to embed the information directly with the image data. Specifically, we recommend the use of Adobe's Extensible Metadata Platform (XMP) [8-10] to store the quality metrics directly in the image headers. XMP is based on XML descriptors, and it is compatible with most common image formats such as TIFF, JPEG, and JPEG 2000. There is currently no standardized XMP schema for quality metadata, and hence we have created one to represent the key quality attributes that are measured from test targets (e.g., sampling, sharpness, noise, etc.). Figure 4 shows an example of a single quality attribute (in this case, sampling as measured by DPI) represented using the Resource Description Framework (RDF) model specified by the XMP standards.

```
<certIQ:Sampling rdf:parseType="Resource"
xmlns:samp="http://www.certifi-media.com/certIQ/1.0/sampling#">
  <samp:Dpi>
    <rdf:Bag>
      <rdf:li>
        <rdf:Description>
          <rdf:value>306,307</rdf:value>
          <samp:Desc>Min,Max</samp:Desc>
          <samp:Unit>DPI</samp:Unit>
          <samp:Direction>H</samp:Direction>
          <certIQ:Src>Target</certIQ:Src>
        </rdf:Description>
      </rdf:li>
      <rdf:li>
        <rdf:Description>
          <rdf:value>305,305</rdf:value>
          <samp:Desc>Min,Max</samp:Desc>
          <samp:Unit>DPI</samp:Unit>
          <samp:Direction>V</samp:Direction>
          <certIQ:Src>Target</certIQ:Src>
        </rdf:Description>
      </rdf:li>
    </rdf:Bag>
  </samp:Dpi>
  </certIQ:Sampling>
```

***Figure 4**. XMP representation of a sampling quality metric (minimum and maximum DPI in the horizontal and vertical directions).*

The extensibility of the XMP approach is very beneficial as new quality measurements and descriptors can be easily added to the metadata. Moreover, the RDF model is a formal framework that facilitates web searches and semantic knowledge activities using the quality metadata .

As can be seen from the example in Figure 4, the downside of an XMP representation is that it is verbose due to the RDF notation. The XMP overhead per image is between 15 - 25 Kbytes with our current schema. As a result, the XMP packet in an image header is best used to store summary quality metrics that encapsulate the essential quality information. Detailed quality information can then be stored in a separate XML file, with a relative path link included in the XMP metadata. This is illustrated in Figure 5.

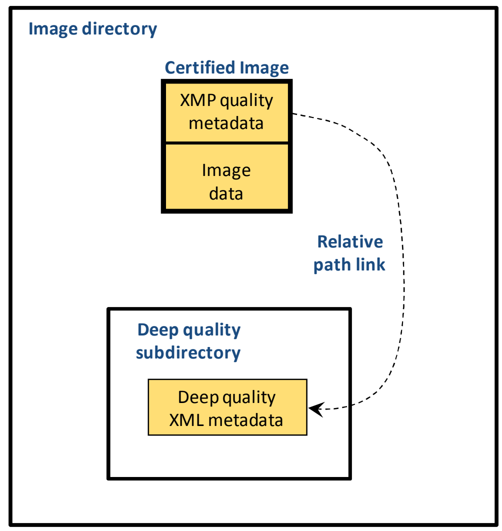

***Figure 5**. Relationship between quality XMP metadata and deep quality XML metadata.*

The XMP framework also makes it straightforward to propagate quality metrics when dealing with intermittent targets in a document queue. As shown in Figure 6, the quality metrics from the last target image can be propagated to subsequent images simply by copying the same XMP quality metadata into the image headers. The XMP metadata can be updated for each image to reflect the number of images or the time since the last target was analyzed.

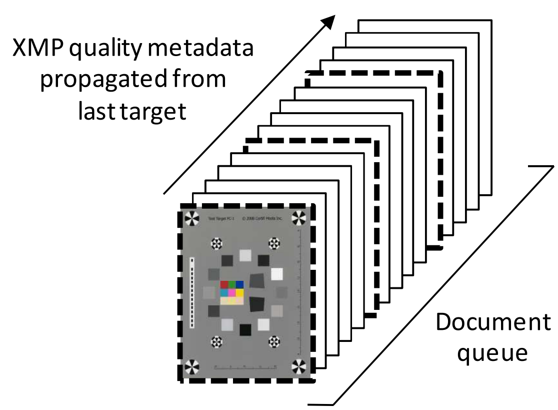

***Figure 6**. Propagation of XMP quality metadata from the last target image to other images in a document queue .*

Security for image data and quality metadata

We recommend the inclusion of a digital signature in the XMP metadata to wrap both the image data and the quality data as illustrated in Figure 7. We refer to an image packaged this way as a "Certified" image. By doing so, the quality metadata is always tied to the image data, which ensures that the quality metrics are always representative of the actual image data and any tampering with either the image data or the quality metadata can be detected.

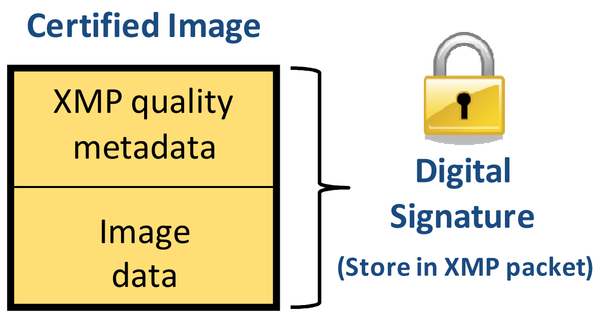

***Figure 7**. Certified image file containing image data, XMP quality metadata, and a digital signature.*

Make exception handling exceptional

An essential part of a QA program is dealing with the quality problems that will inevitably occur. A variety of approaches can be used to simplify exception handling and make it efficient and convenient.

Images that fail one or more quality metrics can be moved automatically to a separate exceptions directory during batch processing/QA. An operator can then review each exception image to judge the severity of the degradation. Images with gross errors, such as blank pages or severe over or under exposure, can be quickly flagged for rescanning. Less severe errors can then be reprocessed to bring the quality metrics within the desired specifications.

Having an interactive tool with integrated QA and image processing as described previously can significantly streamline the reprocessing, particularly when combined with additional extensions of the XMP metadata. Specifically, if the XMP metadata also includes the original image filename and the processing script that was applied to the original image, the Certifi Pedigree Script Builder tool shown in Figure 3 can be used to access a processed Certified image and automatically load the original image and the processing script into the tool. By noting which quality metric is out of spec (sharpness, exposure, etc.), an operator can adjust the corresponding image processing operation until the quality specification is met. This reprocessed image is repackaged as a new Certified file and saved back to the correct filename in the processed image directory to complete the exception handling cycle.

Moving Beyond Fidelity and Target Metrics

In all of the preceding discussion, it was implicit that the quality metrics were produced by analyzing a test target. While target-based fidelity metrics are very useful and robust, there are other types of quality metrics that might be used in a QA program.

One example is demonstrated by referring again to Figure 6, where a test target is used intermittently within a document queue. The issue is quality problems can occur to images between the test targets and it is desirable to trap these problems so exception handling can be facilitated. This scenario points to the need for targetless quality metrics (also called non-reference or blind metrics), which are computed directly from the image data, to supplement the target-based metrics The concept of targetless metrics is not as far fetched as it might sound, and there have been a number of such quality metrics developed in the last several years [11-14]. The extensibility of the XMP quality metadata means that such metrics can be easily included as an adjunct to target-based metrics. It is also possible that targetless metrics could supplant target-based metrics entirely in some applications where it is not feasible to include a test target.

Another example of quality metrics that aren't derived from test targets are metrics produced by automated image processing algorithms. Automated algorithms such as auto deskew (to correct page or text rotation) and auto crop (to crop to page boundaries) will inevitably fail on certain image content. However, these algorithms often include some type of "goodness" measure, which can be used like any other quality metric to trigger exception handling. A simple example is an auto page crop where the typical page size is known and hence the expected dimensions of the output image are known. Cropped images that are larger or smaller than the expected dimensions can be flagged for review.

Examples of quality metrics that aren't based solely on fidelity criteria include metrics that quantify OCR performance. While higher quality generally means a better OCR success rate, strict fidelity measures aren't necessarily required. For example, a tonescale with a higher contrast than normal may yield better OCR results. A metric such as OCR performance is really an integrated measure of several quality attributes such as tonescale, sharpness, noise, uniformity, geometrical distortion, etc. This is an area that remains open to future exploration.

Summary

Moving image quality assurance from being an afterthought or annoyance into the realm of day-to-day practice requires a QA program that:

- Provides meaningful information;
- Is convenient to use; and
- Is cost-effective to implement.

In this paper, we have presented the next steps that are needed to achieve these goals. These steps include fast, robust, and automated analysis tools that are integrated with image processing, combined with an extensible approach to quality metadata handling. We have demonstrated how such goals can already be achieved with today's technology. While the focus was on target-based fidelity metrics, there are clearly new opportunities to move beyond such metrics to include targetless and non-fidelity metrics for certain applications.

References

[1] Federal Agencies Digitization Guidelines Initiative (Still Image Working Group), Digital Conversion – Documents and Guidelines: A Bibliographic Reference (2009). Currently available at http://www.digitizationguidelines.gov/stillimages/documents/Guidelines_Bibliography-2009rev.pdf

[2] Steven Puglia, Jeffery Reed, and Erin Rhoads, Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images, U.S. National Archives and Records Administration (NARA). (2004). Currently available at http://www.archives.gov/preservation/technical/guidelines.pdf

[3] Koninklijke Bibliotheek (National Library of the Netherlands), Metamorfoze Preservation Imaging Guidelines (2007). Currently available at www.metamorfoze.nl/publicaties/richtlijnen/guidelinespijune07.pdf

[4] Hans van Dormolen, Metamorfoze Presrvation Imaging Guidelines, Proc. IS&T Archiving Conf., pp. 162-165 (2008).

[5] Brian W. Keelan, Handbook of Image Quality: Characterization and Prediction (Marcel Dekker, New York, 2002), p. 379.

[6] Michael Stelmach and Don Williams, When Good Scanning Goes Bad: A Case for Enabling Statistical Process Control in Image Digitizing Workflows, Proc. IS&T Archiving Conf., 3, pp. 237-243 (2006).

[7] Don Williams and Peter D. Burns, Measuring and Managing Digital Image Sharpening, Proc. IS&T Archiving Conf.. pp. 89-93 (2008).

[8] Adobe Systems Inc., Extensible Metadata Platform (XMP) Specifications: Part 1, Data and Serialization Models (2008). Currently available at http://www.adobe.com/devnet/xmp/pdfs/XMPSpecificationPart1.pdf

[9] Adobe Systems Inc., Extensible Metadata Platform (XMP) Specifications: Part 2, Standard Schemas (2008). Currently available at http://www.adobe.com/devnet/xmp/pdfs/XMPSpecificationPart2.pdf

[10] Adobe Systems Inc., Extensible Metadata Platform (XMP) Specifications: Part 3, Storage in Files (2008). Currently available at http://www.adobe.com/devnet/xmp/pdfs/XMPSpecificationPart3.pdf

[11] Xin Li, Blind Image Quality Assessment, Proc. IEEE ICIP, 1, pp. 449-452 (2002).

[12] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, A No-Reference Perceptual Blur Metric, Proc. IEEE ICIP, 2, pp. 57-60 (2002).

[13] R. Ferzli and L. J. Karam, Human Visual System Based No-Reference Objective Image Sharpness Metrics, Proc. IEEE ICOP, pp. 2949-2952 (2006).

[14] K. Seshadrinathan and A. C. Bovik, New Vistas in Image and Video Quality Assessment, Proc. SPIE HVEI, 6492 (2007).

Author Biographies

Paul Jones is co-founder and CEO of Certifi Media Inc., which provides automated quality assurance and image processing solutions to improve image digitization workflows. He has 27 years of technical experience in digital image and video processing, image compression, image science, and image quality, and currently holds 29 U.S. patents in these areas. He has previously worked for Eastman Kodak and Xerox. He holds BS and MS degrees in Imaging Science from Rochester Institute of Technology and an MS degree in Electrical Engineering from Rensselaer Polytechnic Institute. He is a member of IS&T and IEEE.

Chris Honsinger is co-founder and CTO of Certifi Media Inc. He has 26 years of technical and leadership experience in digital image, signal, and video processing and in imaging science and security technologies, and currently holds 30 U.S. patents in these areas. He was a co-recipient of the 2005 C.E.K. Mees Award (Kodak's highest research honor) for his work on digital image watermarking. He holds an MS degree in Physics from Ohio University and is a member of IEEE.