Debunking of Specsmanship:

Progress on ISO/TC42 Standards for Digital Capture Imaging Performance

Don Williams
Eastman Kodak Company
Rochester, NY USA

Abstract

For serious imaging practitioners, the benefits of variety and economy in digital capture device selection come as a mixed blessing, among dizzying performance specification hype. This "specsmanship" has created a bazaar-like atmosphere where manufacturers' claims of resolution, speed, and dynamic range resonate like those of so many market barkers. Through regulation, education, and enablement, science-based performance standards vetted through ISO/TC42 will allow many of these claims to be supported or refuted. This paper details the technical content of these efforts and the challenges required to make these standards rugged and easily implemented, as an aid toward less ambiguous and informed device selection.

Introduction

> *"…but he let himself be lured by the siren call of the crowd on First Avenue. It is a common and dangerous mistake of inexperience to get carried away by the treachery of swelling noise."*
>
> Jere Longman
> NY Times

And so it is with capture performance claims of digital imaging devices. One is beckoned by a cacophony of vendor specifications. Loud, confusing (and unregulated), they are, ironically, seductive. The greater the numerical extreme, the greater the allure. For the inexperienced, evaluating these assertions in a consistent, scientific sense is futile. Indeed, even the experienced are disadvantaged without the appropriate tools and guidelines. Let's face it, for the most part, we use these marketing specifications, along with brand name and price, as imaging performance guides. Given the competitive state and relatively low imaging performance expectations of today's consumer digital cameras and scanners, this selection paradigm may actually be reasonable; especially for low demand imaging tasks.

However, for serious amateurs and professionals with demanding projects or clients, relying on this formula as an imaging performance indicator is precarious, especially in the context of high productivity workflow constraints. The difference between sampling frequency (dpi), and true resolution, is confusing. Improved "optical" resolution claims remain suspect. Bit depth alone is far from a sufficient criterion for specifying dynamic range, and the existence of artifacts and noise are dismissed with a shrug. Unlike the world of analog 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 imaging performance standardization efforts of ISO/TC42 are slowly, but surely, changing this free for all. Through the participation of scientists and device manufacturers, a unified architecture of objective signal and noise-based metrics are evolving to help remove device performance ambiguity and robust cross-device comparison. Adapted from proven approaches over a half century of analog imaging experience, these metrics can be used as figures of merit, in their own right, or may be extended as weighted input into higher-order image quality models. A good portion of the standards' practices are the subject of ongoing research as it specifically applies to digital imaging. A compromise between technical rigor and practical execution, they are not perfect but, nonetheless, the best in current thinking.

Of course, the simple issuance of a standard will not ensure its adoption. For this, education, enablement, and improvement efforts are necessary. These have not been forgotten and complement the standard itself by way of classes, technical papers, free software, benchmark testing, and target creation. These are perhaps more important than the documentation itself because they provide practical exercising of the standard by interested users who provide feedback that allows improvements to the standard's practice. These improvements are made through periodic reviews of adopted standards vis-à-vis ISO.

The progress, status, and content of the following TC42 imaging performance standards will be discussed in the indicated groupings. Associated with each standard is an ISO status that ranks, in order, its progression toward full ISO adoption.

I.	Terminology ________________	ISO/DIS 12231
II.	Opto-Electronic Conv. Function__	ISO 14524
III.	Resolution - still picture cameras_	ISO 12233
	Resolution - print scanners ______	ISO/FDIS 16067-1
	Resolution - film scanners ______	ISO/CD 16067-2
IV.	Noise – still picture cameras ____	ISO/FDIS 15739
	Dynamic Range – film scanners__	ISO/CD 21550
IV.	Speed - still picture cameras ____	ISO 12232

WD	– Working Draft
CD	– Committee Draft
DIS	– Draft International Standard
FDIS	– Final Draft International Standard
ISO	– International Standard

Terminology

Frequently forgotten among the techniques and practices outlined in technical standards is the definition or terminology section. While individual standards typically carry their own terminology section, ISO 12231 is a collective document that draws from a number of TC42 electronic imaging working groups (WG 18, JWG 20, JWG 23). As such, it provides a broad perspective on electronic imaging terms. Occasionally, definitions from working groups involved with traditional imaging are also included for completeness. This document should be the first stop for individuals seeking clarification on the meaning of electronic imaging performance terms.

Opto-Electronic Conversion Function- OECF

At the foundation of nearly all of the ISO/TC42 performance standards is the Opto-Electronic Conversion Function (OECF). Similar to a film's characteristic curve that characterizes the transfer of exposure into optical film density, the OECF defines the relationship between exposure, or reflectance, and digital count value of a capture device. By itself, the OECF appears low-tech, but it allows one to evaluate the effective gamma applied to an image, any unusual tonal manipulations, and device non-linearities. Its real power though lies as a rosetta stone for remapping count values back to a common and physical image evaluation space. Without it, meaningful cross-device and cross-parameter performance evaluation would be very difficult. It is the hub to all of TC42's performance standards. This is why it is cited and used so frequently in all of the performance standards.

The single dedicated standard to OECF is for digital cameras, ISO 14524. OECFs use for film and print scanners is nevertheless described and required as defined in the standards' annexes peculiar to those devices. Though the OECF is intimately tied to other performance metrics, its calculation was always made from separate image captures than those metrics of prime interest. This led to inconsistent results between captured frames because of auto-contrast or scene balance algorithms associated with capture devices. For this reason, gray patches for OECF calculation are now being integrated into targets for all of the other performance metrics.

Resolution

" Resolution can serve so many purposes because it does not serve any of them very well"

G.C. Brock, 1968

This observation was made with respect to traditional analog imaging more than a decade before digital imaging began to become popular. Now, sampling and interpolation associated with digital imaging has made the term "resolution" even more ambiguous.

The advertising of device resolution in terms of finished image file size is perhaps the most misleading of all. Through interpolation, an infinite amount of "empty" resolution can be synthetically created that has no physical bearing on spatial detail detection (i.e. real resolution). Short of removing the detector from the camera and physically counting the sensor sites (ugh!) there is no way for the casual user to know the difference. Fortunately, through education, litigation, and standards this practice is becoming less common.

A small but important step towards this is a collaborative draft standard between the Japan Camera Industries Association (JCIA) and International Imaging Industry Association (I3A), "Guidelines for reporting pixel-related specifications". Though not a sanctioned ISO/TC42 effort it bears mentioning because of its relation to device resolution. By strictly defining which sites on a sensor "count" as active imaging sites it removes the "dead" pixel loophole that many manufacturers use to inflate digital still camera pixel count specification. This standard removes the confusion that interpolation techniques themselves impose on resolution, but provides no guarantee of a device's physical ability to provide true spatial resolution.

Yes, simple pixel count (e.g. Mpixels) and sampling frequency (e.g. dpi) are always cited and easy to understand, but Mother Nature frowns at such laziness. She requires that optics, motion, image processing, and electronics contributions also be considered as influencing factors for a device's true resolution. Then and only then is realistic spatial resolution determined. For this, the measurement of Spatial Frequency Response (SFR) or Modulation Transfer Function (MTF) of a device is required. These measurements unify the spatial resolution standards for electronic capture devices under TC42 and are described for cameras (ISO 12233), reflection scanners (ISO 16067-1) and film scanners (ISO/CD 16067-2). Each of these standards adopts a common slanted edge-gradient MTF analysis technique especially suited for digital capture devices. Its accuracy has been benchmarked[1] with both synthetic and real image data. Its chief advantages are ease-of-use, durability, and analytical insight.

The suitability of MTF as an objective tool to characterize spatial imaging performance is well documented and has been used as an image quality prediction tool for more than fifty years[2]. By characterizing contrast loss with respect to spatial frequency, one of its many uses can be to objectively establish the limiting resolution of a device. This is done by determining the spatial frequency associated with a given MTF value, typically 0.1. This frequency is then translated into limiting resolution for a given set of scan conditions and compared to the manufacturers claim to determine compliance. An example of this for a reflection scanner at three different sampling frequencies is shown in Fig 1.

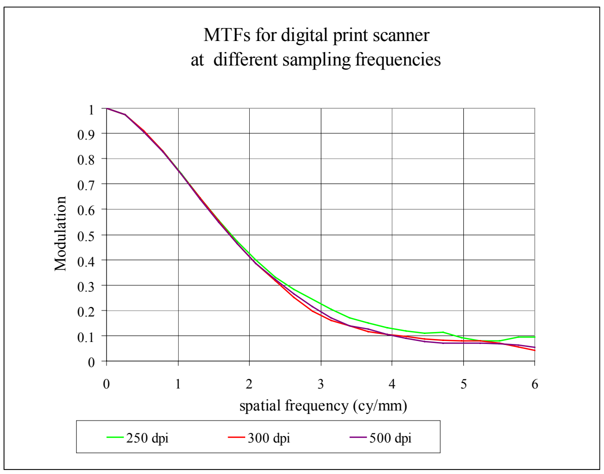

Figure 1

Notice that the MTFs for each sampling frequency (250, 300, and 500 dpi) are essentially identical. The individual curves of Fig. 1 are difficult to identify because they literally overlay. This indicates no real resolution advantage at 300 and 500 dpi compared to the 250 dpi scan. This is indisputable. The 0.1 modulation level corresponds to 4 cycles/mm. Translating this to an effective resolution (dpi = (cycles/mm) * 50.8 ~ 200 dpi), one finds that this scanner is really no better than a 200 dpi scanner, no matter what the advertising claims or sampling frequency. This analysis was performed with tools provided through the TC42/WG18 standards group and is one of many examples where they have been used to objectively clarify resolution performance.

Parenthetically, informative references to ISO 16067-1 detail methods to extract sub-pixel color channel registration errors from the ephemeris MTF data[3]. This artifact is often a problem with linear array scanners and is quantified in the analysis tools provided through the I3A website (www.i3a.org). Color misregistration as large as 1.5 pixels was calculated in the scanner of Fig.1 with the same tools used for MTF calculation.

In the past, MTF measurement has been confined to laboratory settings, and had never matured as a particularly field friendly method for objectively determining resolution; a requirement for widespread adoption and credibility as a standard. This hurdle to acceptance is now largely removed. Through the efforts of TC42 members, free automated software, debugging, affordable high-bandwidth targets, technique documentation, and educational workshops have been provided. The remaining challenges lie in the manufacturing and design of robust targets for film scanners, and improvements to target design for cameras.

Noise & Dynamic Range

"Reproduction quality superior to 4x5 film - no film grain"

actual capture device performance claim

Part of the seduction of digital imaging is the myth that it is noise free. By proclaiming a lack of film grain, the above claim implicitly suggests that this is so. To demystify this, two ISO/TC42 standards are in progress that define noise and dynamic range measurements. ISO/FDIS 15739 is intended for digital still cameras and ISO/CD 21550 for film scanners. Though no effort is currently under way for print scanner characterization of these metrics, the methods and recommended practices of ISO 21550 are likely to apply. The camera standard (15739) is primarily intended to measure noise, but it also makes recommendations on dynamic range. Similarly, the film scanner standard (21550) is primarily intended for dynamic range measurement of which noise characterization is required. Both standards use identical techniques for characterizing dynamic range and noise and are described next.

For the uninitiated, assessing dynamic range in the context of noise may not be obvious. After all, most dynamic range claims are typically tied to device bit depth alone; the higher the number the better. For instance, 12 bits/color (4096 levels/color) would indicate a precision of 1 part in 4096, or a maximum optical density of 3.6 [4] .These simple calculations of dynamic range may be suitable for concept capability tutorials but are far from sufficient for real imaging performance. To understand why, a qualitative definition of dynamic range as applied to imaging applications is needed. I propose the following:

> **Dynamic range** – the extent of energy over which a digital capture device can reliably detect signals: reported as either a normalized ratio (xxx:1) or in equivalent optical density units.

The operative words in this definition are *reliably detect.* Detection is a function of signal strength, the stronger the better. The reliability, or probability, of that detection is a function of the noise associated with that signal, the lower the better. This logic suggests that maximizing the signal-to-noise ratio (SNR) is appropriate for increasing the dynamic range of a device. This was not lost on the members of

TC42/WG18, thus, SNR is integral to dynamic range measurements under the cited standards. They marry signal, (i.e. contrast), with the probability of detecting that signal, noise. So far, so good; we now know *what* to measure. Knowing *how* to measure it is more complex.

Both standards have taken the high road and adopted an incremental SNR approach to the metrology mechanics. Of all the ways to measure signal, incremental signal is probably the most informative for realistic imaging use. Its utility lies in quantifying how well a given object intensity, I_o can be distinguished from another intensity of an arbitrarily small difference, ΔI. Unlike the simple counting of bits, which is a capability measure, this usage of dynamic range is a performance measure as dictated by everyday needs. In the context of a noise it will answer questions like, “Can this capture device distinguish between an optical density of 1.00 and 1.10?” The calculation of the incremental signal is simply the derivative of the OECF function. An example of this for an 8 bit reflection scanner is given in the top of Fig.2

The other portion of dynamic range measurement is device noise characterization. This is determined through a “noise cracking” technique [5] that distills fixed pattern rms noise from random temporal rms noise. Depending on the application, either may be of interest, but in most cases it is the temporal noise alone that is of concern. This step is extremely important for scanners because fixed pattern noise due to the target often accounts for the majority of the total noise. Discounting this target noise is required so that the scanner itself is not discredited. The center graph of Fig. 2 illustrates the noise function.

Taking the ratio of the incremental signal and noise at each OECF patch yields the incremental SNR function. An example of this for a reflection print scanner is illustrated it the bottom of Fig. 2. Dynamic range is then determined from the incremental SNR by noting the density at which a prescribed SNR value is met. For instance, using a typical value of six for a value, the scanner of Fig. 2 would roughly have a dynamic range of 1.5 or 32:1. This measure of dynamic range is significantly lower than the noiseless and flare free capability measure of 2.4 that a simple bit count yields.

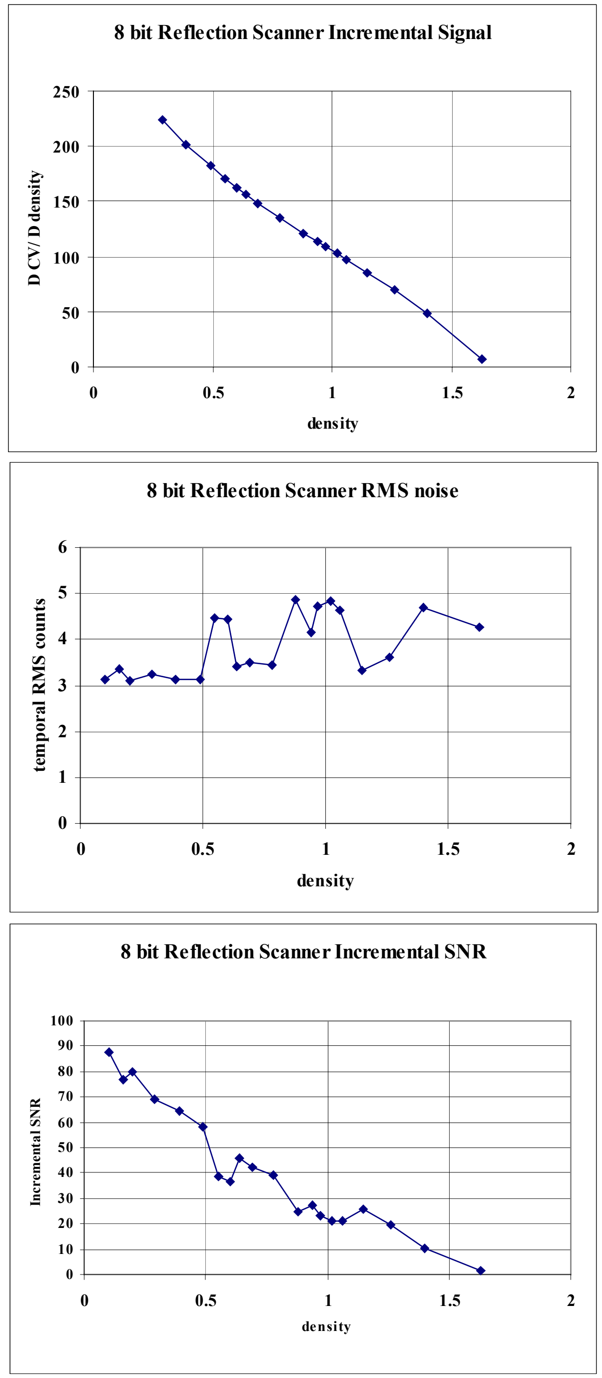

Figure 2

Speed

A camera’s speed rating is the most important attribute in estimating proper exposures for given lighting conditions, and the electronic camera speed ratings in ISO 12232 are meant to match those of their film counterparts, to the extent possible. For instance, using a particular ISO speed value as the exposure index on a digital camera should result in the same camera exposure settings and focal plane exposures as that of a film camera or photographic exposure meter[6]

Currently, almost all digital camera manufacturer’s use this technique for specifying speed. It makes no guarantee on equivalency of image quality only on the exposure required to reach an equivalent mean signal. This “mean signal” criteria for rating speed has served the film community well over the years. It is intuitive and appropriate for film imaging and has withstood the test of time. As much as we try though, formulating an exact equivalency between film and electronic is not always possible. The area of signal amplification is a primary reason.

Perhaps the greatest advantage that digital devices have over film is their flexibility in digitally amplifying mean signal levels to achieve an effective speed. Some flexibility in speed is achievable with film through media and processing selection, but by nature, the amplification step that achieves a given speed point does not fundamentally change that image's utility. Because digital imaging can be so unconstrained in its processing, signals can be amplified or "shaped" to achieve nearly any average output level. This amplifies the noise and often does so to such an extent that it renders the captured image useless. To manage this at low exposures, the preferred method for determining digital camera speed under ISO 12232 considers average output level in addition to noise.

Like dynamic range (ISO 15739), the digital camera speed standard uses an incremental SNR metric. Through image quality studies[7], SNR values of 40 and 10 were chosen to describe excellent and acceptable levels of image quality respectively. The exposure required to achieve these SNR values dictate noise-based speeds. They are referred to as $S_{noise40}$ and $S_{noise10}$ and, like film, are intended to characterize minimum exposure behavior.

Because of film's wide exposure latitude, over-exposure has never been a strong concern. One might pay a small noise penalty for over-exposing but the signal itself was usually not lost. Typically, this is not the case for digital imaging using CCDs or CMOS devices. Ungraceful clipping or blooming can occur in scene highlights for applications where lighting is plentiful, such as studio photography. To accommodate these environments, an ISO saturation based speed has also been defined. This rating helps the user select exposures to prevent these clipping and blooming artifacts and is appropriately used in conjunction with exposure index to help achieve the best possible image quality[6]

For most photo situations, there is little doubt that the noise based speeds are a superior form because of their low-exposure utility and SNR (i.e., image quality) correlation. While it is true that these ratings are not widely cited, a greater interest in doing so has been noted [8]. When this method is widely adopted, the "specsmanship" of speed ratings will certainly be of greater value than it is currently.

Conclusion

The wonderful thing about standards is that
there are so many to choose from.

Anon.

Ultimately, the success of any standard is measured by its level of adoption. And adoption is achieved through regulation, education, and enablement. The regulation of digital imaging performance metrology is well on its way through the vetting and review process provided by ISO/TC42. This paper details the content and status of these standards and includes a view of the scientific rationale for each. With an aim of combining technical rigor and utility, an architecture of sound signal and noise metrology techniques has been established. Though not perfect, the goal is to have them evolve so that one day they may nearly be so.

Education and enablement can accomplish this. Automated easy-to-use software, economical targets, publications, and presentations through I3A and committee members have provided a good start towards this goal. The use of these tools and resources are beginning to allow users to either accept, refute, or at least question manufacturer's "siren calls" in a scientifically sound and unified manner.

Acknowledgments

Gratitudes to Jack Holm, Peter Burns, Sabine Süsstrunk, Ken Parulski, Sean Kelly, and Bruce Pillman for insightful talks on the theoretical and practical aspects of the above standards.

References

1. Don Williams, "Benchmarking of the ISO 12233 Slanted-edge Spatial Frequency Response Plug-in", IS&T PICS Conf., pp. 133-136, (1998).
2. T.H. James, "The Theory of the Photographic Process", pg. 604-606, 1977.
3. Peter Burns & Don Williams, "Using Slanted-edge Analysis for Color Registration Measurement", IS&T PICS Conf., pp. 51-53, (1999).
4. A. Kenney & O. Rieger, " Moving Theory into Practice, Digital Imaging for Libraries and Archives", Research Library Group, pp. 40-41, 2000.
5. Peter Burns & Don Williams, "Distilling Noise Sources for Digital Capture Devices", IS&T PICS Conf., pp. 132-136, (2001).
6. International Standard ISO 12232, "Photography – Electronic still-picture cameras- Speed measurements", 2000.
7. Jack Holm and Sabine Süsstrunk, "An EIQ-subjective image quality correlation study", IS&T's 47th Annual Conference/ICPS 1994, vol. II, pp.634-640.
8. Jack Holm, personal communication.

Biography

Don Williams is a research imaging scientist in the Imaging Science Division of Eastman Kodak Co. and has been with the company since 1981. His primary professional interests lie in rugged signal and noise metrology as they relate to imaging performance. He currently co-leads standards activities ISO 16067-1, ISO 16067-2, and ISO 21550 and frequently contributes and consults on imaging performance for the library and museum communities.