A Pilot Study of Digital Camera Resolution Metrology Protocols Proposed Under ISO 12233, Edition 2

Don Williams, Image Science Associates
Dietmar Wueller, Image Engineering
Kevin Matherson, Hewlett Packard
Hideaka Yoshida, Olympus Imaging Corp.
Paul Hubel, Foveon, Inc.

ABSTRACT

Edition 2 of ISO 12233, Resolution and Spatial Frequency Response (SFR) for Electronic Still Picture Imaging, is likely to offer a choice of techniques for determining spatial resolution for digital cameras different from the initial standard. These choices include 1) the existing slanted-edge gradient SFR protocols but with low contrast features, 2) polar coordinate sine wave SFR technique using a Siemens star element, and 3) visual resolution threshold criteria using a continuous linear spatial frequency bar pattern features. A comparison of these methods will be provided. To establish the level of consistency between the results of these methods, theoretical and laboratory experiments were performed by members of ISO TC42/WG18 committee. Test captures were performed on several consumer and SLR digital cameras using the on-board image processing pipelines. All captures were done in a single session using the same lighting conditions and camera operator. Generally, there was good conformance between methods albeit with some notable differences. Speculation on the reason for these differences and how this can be diagnostic in digital camera evaluation will be offered.

Keywords: Resolution, spatial frequency response, imaging performance, image quality, imaging standards

1. INTRODUCTION

Adopting SFR based protocols for ISO digital camera resolution metrology has proven to be a wise choice, for a number of reasons. One, it allows exercising ones freedom of choice. Rather than constrain the user to single valued committee-defined resolution criterion, it enables one to choose an appropriate SFR response level from the continuous SFR curve to better define task specific thresholds. For instance, depending on the usage case, a 0.50 SFR response level may be a better pass/fail choice than a Rayleigh defined 0.10 level. Simply put, it does as any good standard should; it defines procedures, not specifications. Secondly, SFR is the gateway to enabling image quality measurements of sharpness, or for that case, over-sharpness. Except for good exposure, sharpness is probably the single largest component in customer image quality satisfaction. The literature is literally bursting with citations of frequency response based image quality metrics. SFR is the objective link to such image quality measurements. Finally, SFR can be a diagnostic tool, especially for digital imaging. Non monotonic bumps in the SFR provide signature evidence of spatial image processing operators. When compared across color channels, optical zoom levels, field position, and directions it affords performance insight that single valued resolution measures cannot.

Periodic review of any metrology practice is always good policy. It is especially appropriate to do so now for digital cameras for a couple of reasons. As early adopters of any product can attest (be it software, automobiles, or cool electronic gadget), unexpected behaviors, usually from field environments, always occur. Accommodating for and debugging these in subsequent versions is imperative for continued and credible adoption. Standards procedures are not immune to such flaws either. A number of new proposals in the second addition are intended to fix these flaws.

Secondly, despite the best attempts to forecast future camera processing paradigms and prepare for them, gaps in metrology resilience inevitably occur. For instance, the levels of adaptive and non-linear image processing in today's digital cameras present real challenges in developing metrology protocols that are resistant to such conditioning. Combining or comparing results from differing techniques increases the likelihood of unmasking image processing behaviors and allows the user to make more prudent decisions on camera settings and limits.

For the uninitiated, calculating and interpreting a full SFR curve can be intimidating however. Faced with interpreting an SFR curve, what is the average user suppose to do? Sometimes, all the end user really needs is a single resolution number that will provide a reasonable level of confidence on the camera's resolution performance. A number that does not require one to be among the club of scientists and engineers for whom SFR is best suited: a simple, non-fanciful pass/fail method that is compatible and aligns with deeper level SFR techniques. Because ISO standards should serve a broad level of users, they must speak to different audiences to be accepted and successful. Such a solution, visual resolution, is proposed in the second edition of the standard.

Consistent with the above reasoning then, three significant metrology protocol changes to the first edition of ISO 12233 are being proposed for the second edition of the standard. They are:

1) Low contrast (60% modulation) slanted edge-gradient SFR analysis with improved open source code.
2) Polar coordinate sine wave (Siemens Star) SFR technique with open source software.
3) Visual resolution threshold criteria using 5 line hyperbolic wedge feature with open source software.

A narrative on these three methods as well as the rationale for this pilot study follows.

2. NARRATIVE

Different is not necessarily wrong. It's simply different. The same holds when it comes to resolution measurement of digital cameras. Members of the ISO TC42/WG18 committee on digital camera resolution make no claims on the truth market when it comes to a single best resolution metrology approach. It often depends on the usage case. This is especially so in the chaotic setting of today's adaptive non-linear image processing pipelines that raw image files are required to pass before being delivered to the end user as JPG, TIFF or BMP formatted image files. If there is one truth to digital camera SFR metrology, it is that there is no singular truth. Just like the optical sciences have geometric, physical, and quantum models that explain different phenomenon, different SFR approaches also provide a means to explain increasingly complex and different camera behaviors. By themselves, each can be lacking, but as a group they can provide powerful insight. There are however suggested requirements in order to be considered as a new protocol. Some are:

1) A level of compatibility to previous editions, historical, and companion practices.
2) Significant added value to the current standard,
3) Reliable and resilient field behaviors,
4) A scientific foundation, and
5) Fair and reasonable access to the analysis source code.

Initial individual evidence from each member proponent fulfilled the spirit of these items. But, no unified experiment exercising each protocol had been performed under controlled laboratory conditions with all advocates present. That was the purpose of this pilot study: to openly exercise the protocols of each method under equivalent and agreed upon capture conditions. This left little room for debate on experimental bias since all interested members participated. These three methods are briefly described below.

The original SFR metrology method, edge-based spatial frequency response (E-SFR), is identical to that described in the first edition, except that a lower contrast edge is recommended for the test chart. Regions of interest (ROI) near slanted vertical and horizontal edges are digitized and used to compute the SFR values. The use of a slanted edge allows the edge gradient to be measured at many phases relative to the image sensor photo elements, to yield a phase averaged SFR response. It is most consistent with ISO 15529 which uses a line feature instead of an edge.

A second sine wave based SFR (S-SFR) technique is introduced in the second edition. Using a sine wave modulated starburst target, it is intended to yield an SFR response that moderates spatial frequency signatures introduced by aggressive non-linear image content driven processing of many digital cameras. In this sense, it is meant to enable easier interpretation of SFR responses from such camera sources. Comparing the results of the edge-based SFR and the sine based SFR may indicate the extent to which non-linear processing is used.

While resolution and SFR are related metrics, their difference lies in their comprehensiveness and utility. A proposal for the second edition of ISO 12233 is a single frequency parameter, visual resolution, which indicates whether the output signal meets a minimum threshold criterion of detail information for visual detection. It can be very valuable for rapid manufacturing testing, quality control monitoring, or for providing a simple metric that can be easily understood by end users. The algorithm used to determine resolution has been tested with visual experiments using human observers and correlates well with their estimation of high frequency detail loss. Early feedback suggests that it correlates well with the spatial frequency at the 0.10 SFR response level.

It is emphasized that this is a pilot study. It is not meant to be comprehensive but rather a “sniff test” to gauge the suitability of the committee draft (CD) document to proceed to the Draft International Standard (DIS) level. Equally important, it is also meant to provoke thought on more comprehensive methods for future experiments vetting these techniques.

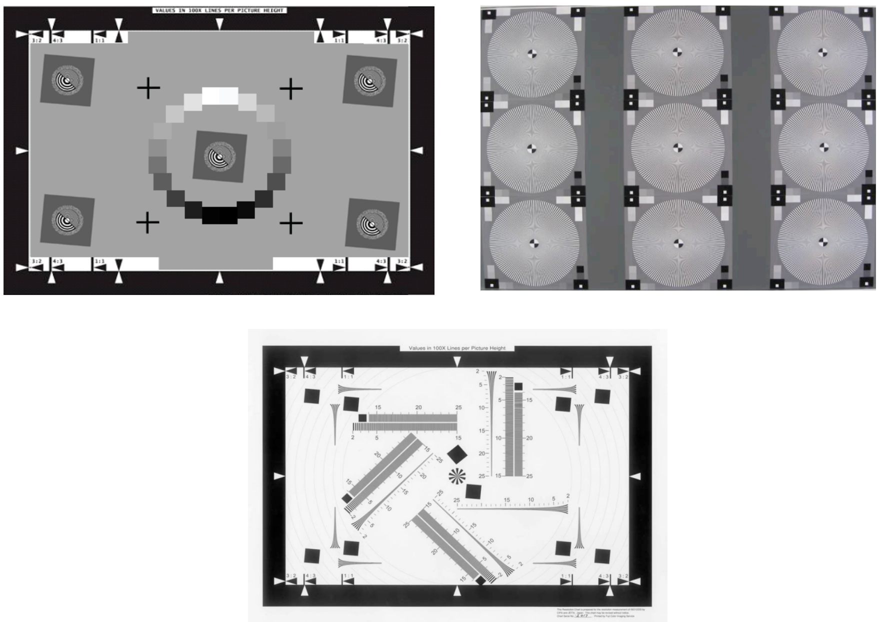

Fig. 1 – Candidate Target formats supporting the three proposed methodologies of the ISO 12233, second edition. a) Low contrast slanted edge-gradient (upper left), b) Polar sine-wave, Siemens star (upper right), c) Visual resolution.(middle)

3. EXPERIMENTAL

All image captures were done under identical studio conditions (vertical target mounting, lighting, camera-to-target distance, etc.) with the same trained camera operator, and session at Image Engineering near Köln, Germany. The target features of interest were all on the optical axis and covered roughly the center one-tenth of the field of view. Ideally, one target that included all suggested feature sets should have been used. This would prevent any frame to frame variability introduced by the camera processing. It would also rule out any variability due to focus. Lack of experimental foresight prevented this though and is one item that should be remedied in future testing. Eight different camera models in the 3.0 to 6.0 Megapixel range from 4 differently labeled brands were used. All had Bayer CFA patterns with the exception of one.

The charts were illuminated using D50 illumination. The uniformity of the illumination over the area of the chart was 5% and the illumination level was EV 7. During image capture, the chart was oriented parallel to the focal plane of the camera under test, and its horizontal edge is parallel to the horizontal frame. The camera distance to the target was adjusted slightly to frame each target in a 4:3 aspect ratio. The performance of each algorithm was tested with the following camera settings: ISO 100, auto white balance, low, medium, and high sharpening, camera RAW mode, and camera JPG best. Camera strobe was turned off.

The experimental workflow was as follows. After agreeing upon setup conditions (i.e., zoom level, processing settings, and file format) for a given camera, the camera was tripod mounted, and auto-focused on the area of interest for the E-SFR target. A frame was captured by manually and gently depressing the shutter release. Without refocusing this was repeated nine more times for a total of ten replicate frames for each condition. The E-SFR target was removed and replaced by a similar sized S-SFR target. With no camera changes made, the camera was manually auto-focused and ten replicate frames captured as before. Finally the visual resolution target replaced the S-SFR target and the process repeated. Another camera was then mounted and the entire above process was repeated for that camera and its settings.

Each of the ten replicate captures per camera was analyzed and the representative ones used for the analysis and discussion. Renderings of the target types used are illustrated in Fig. 1.

4. RESULTS AND DISCUSSION

After reviewing the SFR analyses and visual resolution results of the image captures from each camera, representative SFR curves were chosen for discussion below. All results were derived from on-axis target features in both the horizontal and vertical directions. The slanted edge gradient and the Siemens star targets have features in each direction that enable replicate SFR estimates in the horizontal and vertical direction. Two replicate estimates for each direction and each technique were performed and are plotted individually. The slanted edge analyses were performed using the Matlab® software provided on the International Imaging Industry Association (I3A) website (http://www.i3a.org/downloads_iso_tools.html). These results are indicated by the dashed line curves and are labeled as “E-SFR”. A user guide and a document on the E-SFR calculation logic can also be found at this site.

The Siemens start SFR results were estimated from Matlab® software provided by Image Engineering (http://digitalkamera.image-engineering.de/index.php/Downloads). These are labeled in the below graphs as “S-SFR” and are indicated by the solid black line in the plots. A user guide, in addition to comprehensive details, on the S-SFR calculation logic can be found at the Image Engineering website.

Visual resolution results were estimated using the HYRes software provided by the Camera and Imaging Products Association (CIPA) (http://www.cipa.jp/english/hyoujunka/kikaku/cipa_e_kikaku_list.html#). A detailed document on the software logic as well as a user guide can be found there also. Since visual resolution is a single valued metric, the resolution estimates from this method are indicated on the below graphs as a single vertical line at the appropriately calculated frequency. This frequency is labeled as “HYRes”. Since replicate features in each direction were not available on the target, only one estimate per direction was calculated. It is worthwhile noting that the visual resolution calculations for this experiment were done on a nine line hyperbolic resolution wedge. The proposed target in the second edition of the standard however calls for a five line resolution wedge. Given the single line rejection mode of the HYRes software logic, it is likely that HYRes results may be biased to higher frequencies using the lower number line feature of the new proposed target.

Before discussing selected individual results below, broad comments on the results are in order. There was good conformance and stability between the three methods, albeit with some notable exceptions. Previous anecdotal claims of large differences in the SFR estimates were generally not observed and visual resolution results aligned well with 0.10 SFR response levels for the most part. The camera results chosen below for discussion are intended to demonstrate the gamut of good and bad conformance between the methods. It is again emphasized that this experiment was not intended to be comprehensive but rather a nominal gauge of suitability for inclusion into the next edition of the ISO standard.

4.1 Camera A – Default RAW and nominal unsharpened JPG settings

The reader is referred to Fig. 2 below. Considering the vastly different approaches of the E-SFR and S-SFR methods as well as an experimental environment, the SFRs derived from both RAW and processed JPG images of Camera A were encouraging. While full SFR conformance for different file types was unexpected because of image processing differences, reasonable SFR agreement was anticipated between the proposed methodologies. This is demonstrated in the graphs of Fig. 2. Though there was a consistent low bias of the E-SFR results relative to those of the S-SFR it is not considered dramatic. Based on the results of a companion camera product, Camera C, there is reason to suspect that these differences are real and a result of differential image processing of edges features. There was also a consistent difference in E-SFR estimates between replicate directional edges: as if there were clear processing change between light-to-dark transitions compared to dark-to-light transitions. The HYRes results for visual resolution tracked well with the 10% response level of each SFR's estimates.

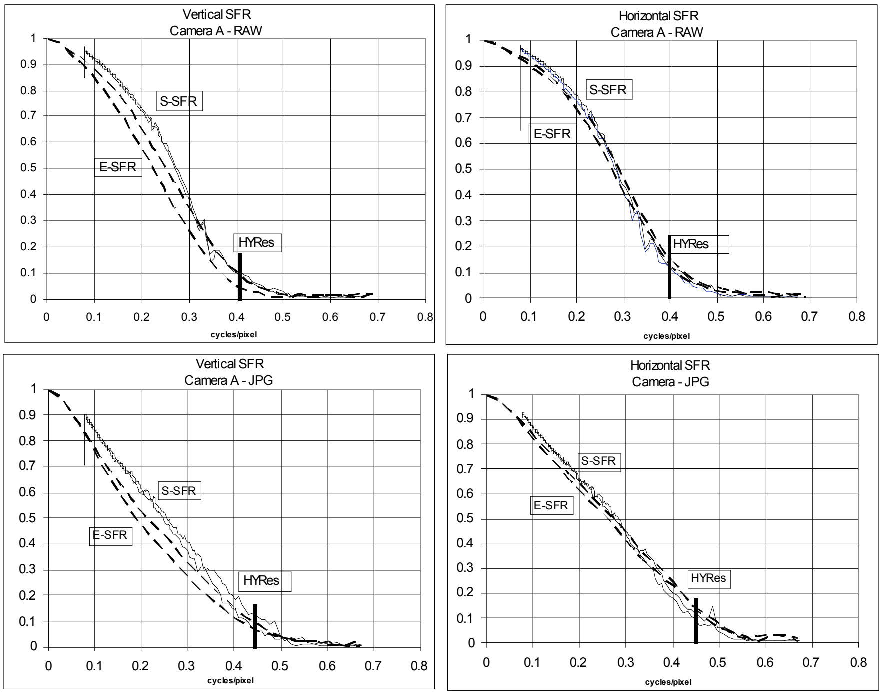

Fig. 2 – Results for Camera A

4.2 Camera B – Default sharpened JPG settings

These results, Fig. 3, too were encouraging. There was little doubt going into this experiment that there would be differences between the SFR methods. This set of results demonstrated that these results were indeed different at exactly the sharpened frequency bands expected while still yielding excellent tracking at the higher frequencies. The sharpening processing was aggressive enough to affect the S-SFR estimates also. Such differences between the E-SFR and S-SFR estimates could be suitable indicators of extracting information on sharpening algorithms.

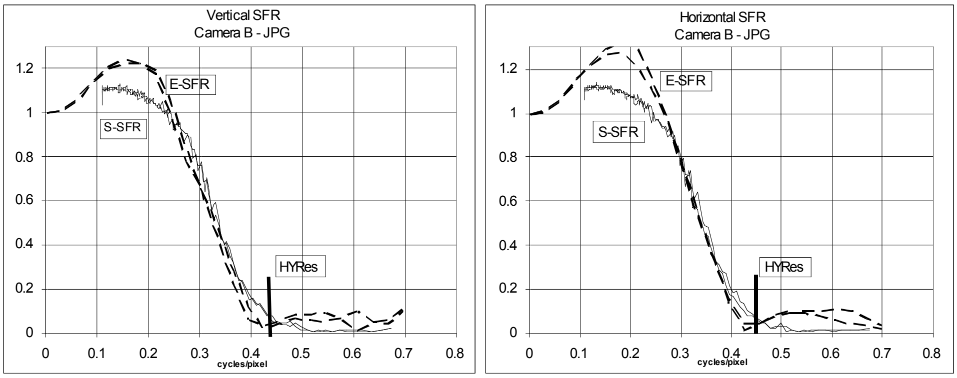

Fig. 3 – Results for Camera B

4.3 Camera C – Low and high sharpening levels for same camera

The results for this camera, Fig. 4, demonstrate the level of diagnostic potential in using the different SFR methods in comparison. There was generally good compliance between to two SFR approaches, on average, but there were remarkably extreme behaviors that required investigation. There was the question of why the one set of E-SFR estimates were consistently biased low for the low sharpening setting. The high sharpening condition also revealed significant differences between the two replicate directional edges. On average, the E-SFR numbers matched the S-SFR results but the component estimates in that average were very different. The edge region-of-interest (ROI) associated with the low biased behaviors were left-to-right (horizontal component) and top-to-bottom (vertical component) light-to-dark edge transitions respectively. Such behaviors required detailed look at the images. Upon inspection there were clear edge sharpening differences between the same directional edges depending on whether the edge transition occurred from light-to-dark compared to dark-to-light. This was felt to be an image processing imposed artifact rather than an optical one. Tracking such differences between methods can be a potentially powerful diagnostic tool for image performance management and the choice of SFR methods.

4.4 Camera D – RAW and processed TIFF images from same camera: no sharpening

The differences between E-SFR and S-SFR methods in Fig. 5 were quite distinct. In all cases the S-SFR gave greater estimates than the E-SFR. Perhaps the problem lay with the significant power beyond the half-sampling frequency. At this writing, the reasons are inexplicable. One of the goals of this study was to illicit these type behaviors for future investigation. To that extent, the experiment was a success. It should be noted that the HYRes result for visual resolution did a very good job in limiting the maximum spatial frequency to the half sampling frequency.

4.5 Epilogue

A common, and natural, question is often asked about differing SFR techniques, “Which is most accurate?” The answer is very simply that they are equally accurate, depending on the signal content of interest. Though not demonstrated here, both proposed SFR techniques have been shown to be equally accurate when tested with synthetically created edges (E-SFR) and sine waves (S-SFR). It is largely the signal dependent nature of today’s digital camera processing that introduces the differences shown here.

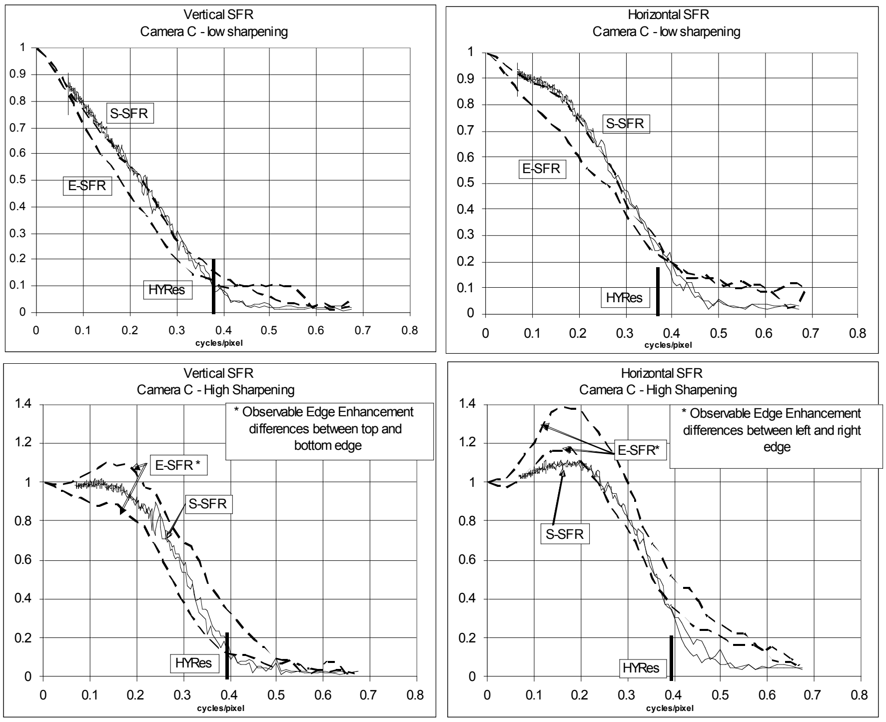

Fig. 4 – Results for Camera C

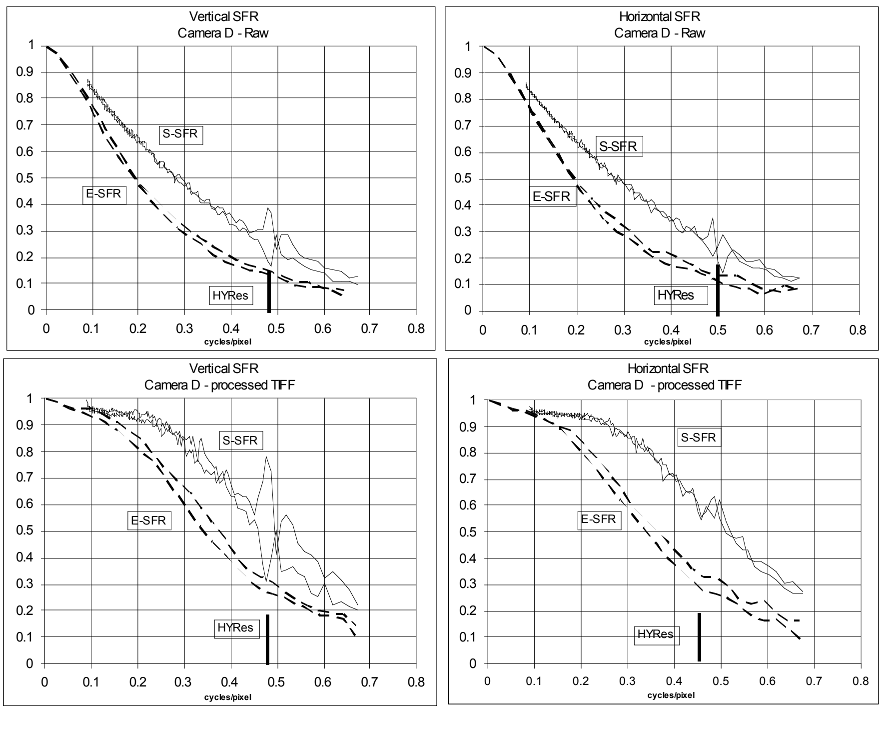

Fig. 5 – Results for Camera D

CONCLUSIONS

The results of this pilot study are very encouraging and are testimony to the benefit of cooperative agreement and controlled experimental technique when doing such comparative studies. The anecdotal claims of large differences between SFR methods did not prove true. While there were some exceptional differences between some camera's results, there were also clear explanations for these differences. These variations could complement the SFR analysis process by providing comparative SFR signatures to diagnose camera behaviors. Future experimental design should include replicate captures over multiple sessions as well as the use of a single target that includes all pertinent resolution features with the same contrast levels. This latter item would help eliminate contrast dependent image processing and operator variability as the cause of differing results between metrology techniques.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the helpful comments and suggestions from the members of the ISO 12233 standards revision committee.