

Estimation of ISO12233 Edge Spatial Frequency Response from Natural Scene Derived Step-Edge Data

Oliver van Zwanenberg[^]

PhD Research Student, School of Computer Science and Engineering, The University of Westminster, UK
E-mail: o.vanzwanenberg@my.westminster.ac.uk

Sophie Triantaphillidou[^]

Professor, School of Computer Science and Engineering, The University of Westminster, UK

Robin B. Jenkin[^]

Principal Image Quality Engineer, NVIDIA Corporation, USA
Visiting Professor, School of Computer Science and Engineering, The University of Westminster, UK

Alexandra Psarrou

Reader, School of Computer Science and Engineering, The University of Westminster, UK

Abstract. *The Natural Scene derived Spatial Frequency Response (NS-SFR) is a novel camera system performance measure that derives SFRs directly from images of natural scenes and processes them using ISO12233 edge-based SFR (e-SFR) algorithm. NS-SFR is a function of both camera system performance and scene content. It is measured directly from captured scenes, thus eliminating the use of test charts and strict laboratory conditions. The effective system e-SFR can be subsequently estimated from NS-SFRs using statistical analysis and a diverse dataset of scenes. This paper first presents the NS-SFR measuring framework, which locates, isolates, and verifies suitable step-edges from captures of natural scenes. It then details a process for identifying the most likely NS-SFRs for deriving the camera system e-SFR. The resulting estimates are comparable to standard e-SFRs derived from test chart inputs, making the proposed method a viable alternative to the ISO technique, with potential for real-time camera system performance measurements. © 2021 Society for Imaging Science and Technology.*

[DOI: 10.2352/J.ImagingSci.Technol.2021.65.6.060402]

1. INTRODUCTION

The ISO12233 slanted e-SFR is a standardized digital camera system performance measure [1], e-SFRs are measured from slanted step-edges selected from images of suitable test charts, captured under controlled laboratory conditions.

A perfect edge contains infinite spatial frequencies at a constant modulation. The system degrades edges and the Modulation Transfer Function (MTF), or its sister measure, the SFR [2], is measured using spatial frequency analysis of this degradation. Practically, SFRs can be derived from

any input step-edge, as long as the edge is characterized and calibrated in the output measurement.

Test chart step-edge(s) used for measuring e-SFRs according to ISO12233 are isolated within regions of interest (ROI). For each ROI, the edge is located to sub-pixel accuracy, and the position of each pixel in each row is projected down the slope to produce a super-sampled Edge Spread Function (ESF). The ESF is subsequently differentiated to obtain the Line Spread Function (LSF) [3], and the modulus of the Fourier transform is calculated to yield the e-SFR.

The research presented in this paper aims to take the standard e-SFR algorithm away from laboratory settings and the use of test charts, and applying it to natural scene inputs to obtain an e-SFR comparable to the ISO method. Natural scene derived camera system e-SFR measurements, amongst other advantages, open possibilities to real-time SFR monitoring, suitable for applications related to intelligent vision systems.

This paper presents the complete framework for obtaining NS-SFRs and subsequently uses this data to estimate the system e-SFR. Firstly, a background is provided for scene derived camera performance measures. The methodology for extracting NS-SFR data from a given camera dataset is described. Subsequently, the variation introduced to the results with respect to input edge and ROI parameters is studied. The study provides insights for setting parameter ranges to obtain stable system e-SFR estimates. The methodology for estimating the system e-SFR is discussed, followed by an overview of estimate validation; a complete analysis of the results is presented in [4] by the same authors. The paper concludes by discussing advantages and improvements to the current version of the framework and further work required for real-time implementation.

[^] IS&T Members.

Received Aug. 6, 2021; accepted for publication Oct. 24, 2021; published online Nov. 18, 2021. Associate Editor: Dmitri Anatolyevich Gusev.

1062-3701/2021/65(6)/060402/16/\$25.00

2. NATURAL SCENE DERIVED CAMERA PERFORMANCE MEASURES

Measuring SFR from natural scenes is not a new concept. Previous studies have used natural scene step edges to assess aerial imagery [5] and optimizing digital scan resolution for film archives [6]. Both these approaches presented robust methods for using natural scene step-edges as a qualitative assessment. Neither were further developed to acquire reliable performance measures from natural scenes compatible with the e-SFR standard measure.

Beyond edge-based natural scene measurements, the texture-MTF method was employed, using test charts such as dead-leaves [7, 8]. Texture-MTFs from images of scenes provide signal-and-process-dependent MTFs (SPD-MTFs) [9–11]. SPD-MTFs were proposed for assessing the impact of non-linear image signal processing (ISP). Results can be used to improve the performance of relevant signal-to-noise image quality metrics (IQMs) [11, 12]. The texture-MTF method traditionally uses test charts with a known input signal and noise spectra, such as dead-leaves [7] or spilled coins [13]. In natural scenes, both are unknown. To resolve this problem, the first study in this area used prints of natural scenes as test charts and high-resolution scans to obtain them [9]. The second method used camera pipeline simulations with known signal and noise spectra variations [11]. In real-world applications, the natural scene signal and noise spectra are inaccessible; the texture-MTF is therefore not appropriate.

Another way is to measure the camera system MTF directly from the Point Spread Function (PSF) [3]. The PSF is rarely used as it is impractical to obtain an impulse of light narrow enough. Nonetheless, the most advanced method to date to obtain camera MTFs from natural scenes uses a convolutional neural network (CNN) to estimate the PSF [14]. The CNN was trained using linear camera pipelines characterized using a PSF panel, i.e., a panel of black illuminated pinholes, to obtain the ground truth PSFs. The outputs are MTF estimates with high accuracy across several Digital Single Lens Reflex (DSLR) systems [14].

Advantages include the MTF being estimated across the entire frame and relatively low computation time, at a few minutes per image [14]; it is not a real-time measure but provides MTF data quickly compared to obtaining it under controlled laboratory conditions. Although, a single image can be used for measurement, multiple images were shown to improve accuracy. The method's drawbacks are that it is not compatible with a standardized method, and importantly, the ground truth is difficult to obtain (needs specialist equipment). The CNN is currently trained using several linear systems; there is no information reported on non-linear ISPs and whether an SPD-MTF can be measured.

The method proposed in this work was introduced in three brief publications. The first set the basis for measuring SFRs from natural scenes and examined the best-suited edge detection technique for the purpose [15]. The second described and initially tested a preliminary version of the NS-SFR framework [16], and the last provided an overview of system e-SFR estimation [17]. This paper expands on these

publications, detailing a refined methodology for measuring NS-SFRs from natural scene image datasets and subsequently using this data to derive the system e-SFR.

It is important to note that the NS-SFR is not considered a camera system performance measure since it is derived from uncharacterized scene edges with unknown spatial frequency content. Instead, it is a function of both the system performance and the scene content. The system e-SFR is the system performance component of the NS-SFRs. To derive the system e-SFR, ideally, the effect of the system and that of the scene contents should be separated. This separation is not straightforward without either being characterized. Statistical analysis of the NS-SFRs to define the most stable edge contents corresponding to perfect step-edge inputs was pursued to estimate the system e-SFR.

There is a drawback when obtaining the system performance from scene inputs, affecting all methods discussed in this paper. If the power spectral distribution of the image does not contain sufficient content to employ a specific performance measure, it cannot be implemented. For example, in context to the proposed NS-SFRs, if there are no step-edges containing appropriate spatial frequencies in the scene, a system e-SFR cannot be derived.

3. NS-SFR FRAMEWORK

3.1 NS-SFR Framework Overview

The proposed framework comprises a series of processes used to locate, isolate, and verify step-edge ROIs from pictorial natural scenes. Edges in the selected ROIs are then used to measure NS-SFRs. The framework is implemented in a sequence of images (i.e., an image dataset), originating from the camera system to be characterized, as summarized below:

- (1) Edges in the captured scenes are located using an adapted Canny edge detector [18].
- (2) Edges that exceed 5 pixels from neighboring edges are cropped, forming ROIs.
- (3) The edge of interest within each ROI is isolated using pixel stretching [16]. This procedure removes texture, noise, and any undetected edges from the ROI.
- (4) The gradient across each ROI is used to validate the appropriate step-edge profile.
- (5) These natural-scene ROIs are processed through the traditional slanted-edge algorithm, producing a series (an *envelope*) of NS-SFRs per each image. The *sfrmat4* [19] is currently utilized for this purpose.

Figure 1 illustrates this procedure.

3.2 Edge Detection

Canny edge detection [18] was assessed against a matched detection method using normalized 2D cross-correlation [20, 21]. The Canny detector was deemed the most appropriate for this purpose [15], but other gradient detection methods can be used. The Canny detector has three advantages. First, its dual threshold provides edge

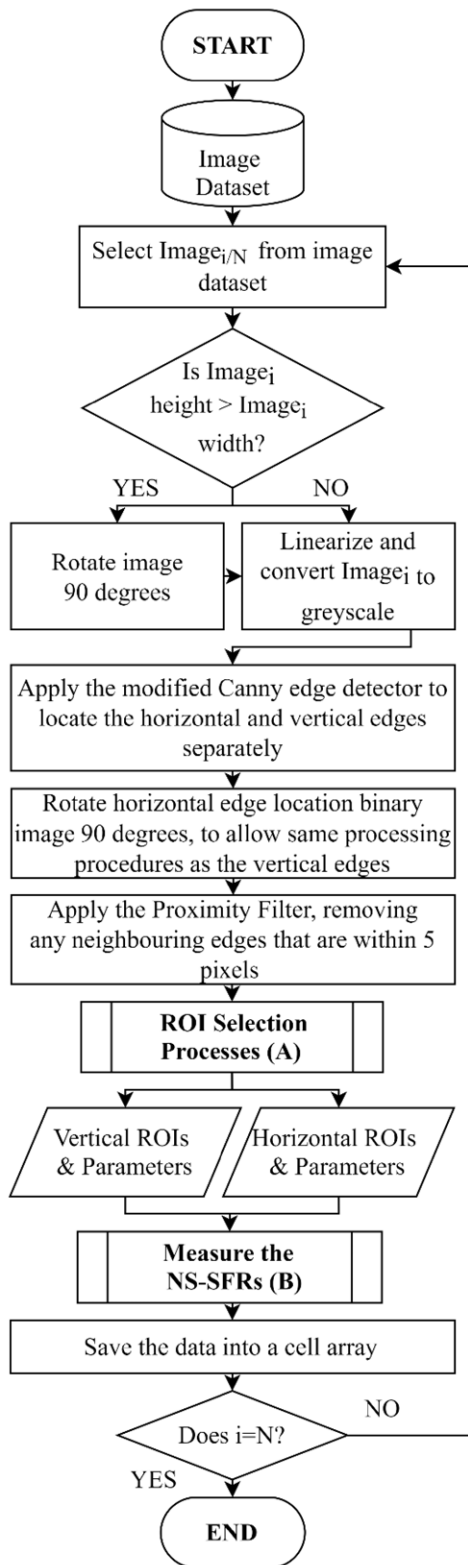


Figure 1. The NS-SFR framework used to measure the NS-SFRs from a natural scene image dataset.

locations less degraded by image noise. Second, these thresholds can be obtained automatically, adapting to the image content through the Otsu method [22]. Finally, the Canny algorithm was adapted to keep the horizontal

and vertical edge orientations separate by not performing the magnitude calculation. Edge orientations need to be segmented since the system e-SFR is dependent on the edge orientation.

3.3 Region of Interest Selection

After identifying ROI edge locations, edges that are in close proximity are removed. These edges are usually found in areas of high busyness. When neighboring edges are too close, their ESFs overlap and therefore, these edges cannot be processed by the e-SFR algorithm. A proximity filter [16] is used to remove edges within 5 pixels of a neighboring edge. A minimum of 5 pixels between edges is used to allow a single edge of interest to be isolated.

The refined binary image with the edge locations for horizontal and vertical edges, and linearized greyscale image are passed through an ROI selection process, as illustrated in Figure 2. This selection process consists of cropping the potential ROIs, utilizing the MATLAB Bounding Box function [23].

All resulting bounding box coordinates are adjusted to give an additional 5 pixels on either side of the ROI, ensuring that the edge is not positioned in the corners of the ROI window. ROI widths narrower than 20 pixels were expanded to meet the minimum width necessary for e-SFR processing. This threshold is derived from data gathered in section 4.2, reducing the error in the NS-SFR. In addition, the boundary height coordinates are set so that the minimum crop height is also 20 pixels and the maximum is 128 pixels. ROIs with a height under the lower threshold are deselected. Crops exceeding 128 pixels are segmented into smaller ROIs, prioritizing the selection of ROIs with a height of 128 pixels before cropping smaller ROIs.

Ideally, small ROIs should be used to reduce the probability of including unwanted artefacts, including changes in illumination, focus shift across the edge, double-edges and overlapping scene structures. Also, smaller ROI heights were permitted to expand the amount of data from each scene. Thus, the maximum height is still relatively small compared to the ISO standard, but as shown in section 4, the ROI window sizes introduce only slight SFR variations thanks to the framework isolation techniques.

The rationale for using an adaptive width, keeping the ROI as narrow as possible without introducing error, and prioritizing the height selection of 128 pixels, was to minimize noise through a high vertical-to-horizontal aspect ratio [24], as detailed in section 4.

3.4 Pixel Stretching

Once the potential ROIs are cropped, the edge of interest is isolated, ensuring removal of any unwanted scene content from the ROI. This isolation is achieved by developing pixel stretching [16], influenced by the filtered tails procedure [25] that blurs either side of a noisy edge.

Before applying pixel stretching, the ESF mask is obtained. The ESF mask is the area of the ROI where the edge and its distribution are located. This mask is obtained by

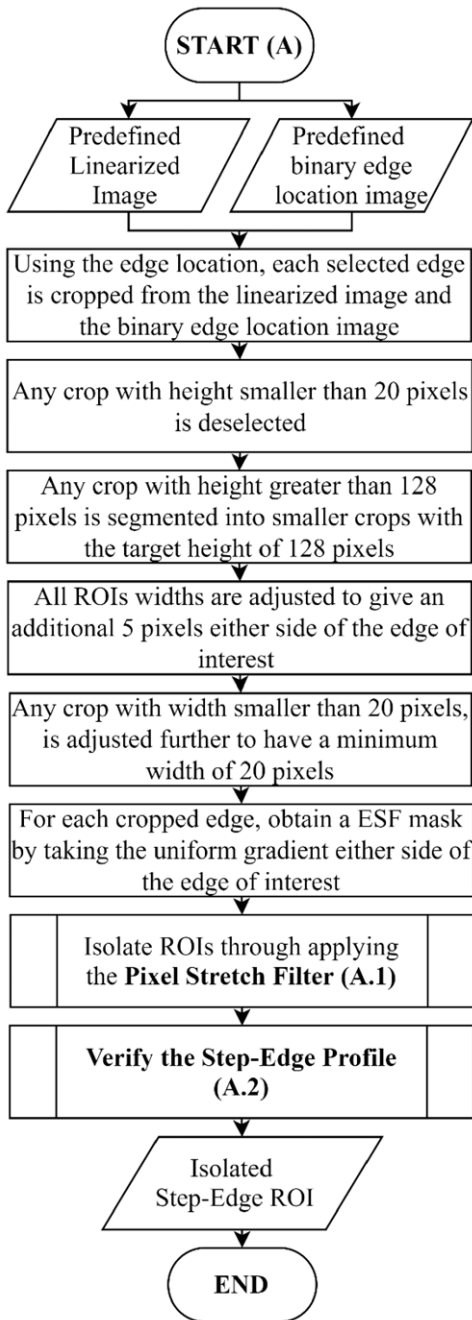


Figure 2. The ROI selection process flowchart outlines the key steps to crop edges, isolate them, and verify they have the required step-edge profile. This process is marked with ‘A’ in the framework in Fig. 1.

finding where the gradient on either side of the edge location becomes zero, setting a threshold to account for the noise floor. This threshold is set at a normalized value of 0.02. Furthermore, a three-pixel dilation is applied to expand the ESF mask border. This is because ESFs commonly have a gradual reduction in the gradient before becoming zero; this area can be cut short by the noise floor threshold. Thus, this dilation ensures that the ESF is within this masked area.

The ROI crop and corresponding ESF mask are used as the input to pixel stretching. For each pixel row in the

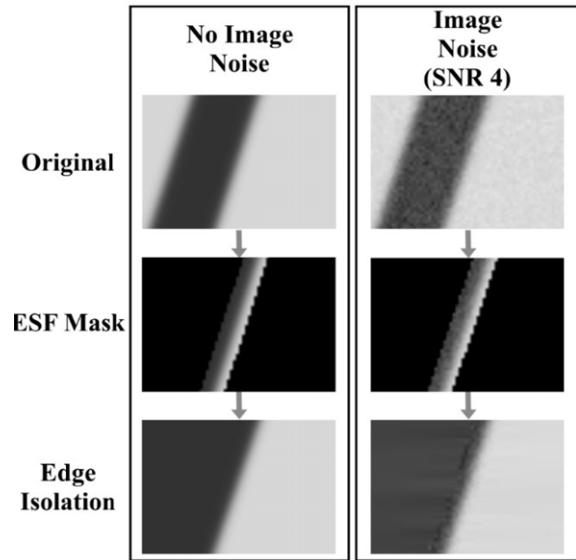


Figure 3. A central ROI crop segment of a simulated step-edge without image noise and at SNR4.

ROI, coordinates for the left (x_L, y_L) and right (x_R, y_R) mask boundaries are located. Two ‘T’-shaped median values are calculated. The left value is computed from the coordinates (x_L, y), ($x_L - 1, y$), ($x_L - 1, y - 1$) and ($x_L - 1, y + 1$); and the right from (x_R, y), ($x_R + 1, y$), ($x_R + 1, y - 1$) and ($x_R + 1, y + 1$). If the mask is at either side of the ROI width boundaries, only one median is calculated. These values are then used to fill the entire row on either side of the mask.

A method such as this can cause a streaking artifact. Although the ‘T’-shaped coordinates help reduce it due to sampling across multiple rows, streaking may still occur under extreme noise levels. Therefore, a weighted Gaussian blur is applied, with the blur weight decreasing toward the boundaries of the ESF mask, whilst the mask remains untouched to maintain the edge of interest.

The ESF mask and pixel stretching isolation are illustrated in Figure 3. The figure shows how in ROIs without noise, the edge can be isolated effectively. ROIs with a low signal to noise ratio (SNR) can isolate the edge with minimal streaking.

The flowchart in Figure 4 lays out the processes used to pixel stretch a ROI.

Besides significantly increasing the number of edges isolated from natural scenes, pixel stretching has the benefit of reducing noise and non-uniformities.

3.4.1 Noise Reduction

The noise reduction works in a similar fashion to the filtered tails procedure [25], removing unwanted variance on either side of the edge. Figure 5 illustrates measuring e-SFR from a noisy edge with and without pixel stretching. Noise was simulated using Poisson and Gaussian distributions to replicate shot and read noise, respectively.

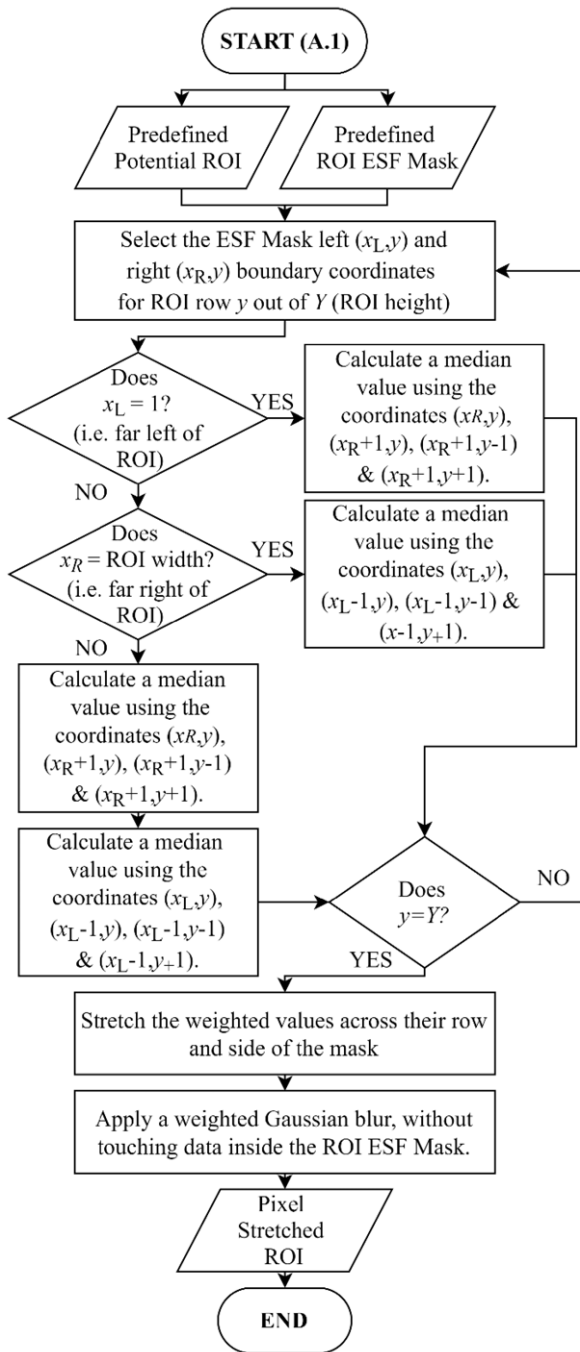


Figure 4. The pixel stretching flowchart details how edge isolation is achieved. This process is marked with ‘A.1’ in the framework in Fig. 2.

3.4.2 Non-uniformity Reduction

Uneven edge illumination causes error in the SFR due to the low-gradient ESF value changes [26]. It is particularly an issue in extracting edges from uncontrolled natural scenes.

Pixel stretching reduces the effects of non-uniformity, as shown in Figure 6. The non-uniform illumination was simulated using a low gradient overlay. Although the improvement is significant, allowing such edges to be used, it does not entirely remove the bias introduced.

3.5 Step-Edge Verification

After edges of interest are cropped and isolated, the ROIs pass through a verification process to remove any that do not contain suitable step-edge profiles.

The step-edge verification flowchart is given in Figure 7. A gradient is measured for every row in the ROI, checking to see if a single positive or negative gradient is present exceeding a noise floor threshold. This process is illustrated in Figure 8. The threshold is again set at 0.02.

The percentage of the total number of rows positively labelled as a step-edge is calculated. This threshold is set at 50% of the ROI height; when above this threshold, the ROI is kept, otherwise, it is deselected.

These gradient and percentage thresholds are set to balance obtaining suitable step-edges and deselecting incorrect edge profiles. The low gradient threshold provides a sensitive process to gradient changes, including image noise and low contrast scene content. As noise is random, the rows labelled as non-step-edges will be randomly distributed through the ROI, whilst non-step-edge rows due to scene content will be in concentrated areas of the ROI.

The 50% threshold provides a counterbalance to the sensitive gradient threshold. Intersecting low contrast scene content would be detected and the ROI removed, whilst the percentage threshold would allow ROIs with some noise to pass. ROIs with noise exceeding 50% of the rows is not ideal and are deselected. These thresholds were determined using an empirical examination from datasets described in section 5.2. Ideally, these thresholds should adapt to the noise floor of the image rather than being static values, preventing noisy edges from being deselected, as long as the edge contrast is also above the noise floor.

This current solution is not ideal, as it allows some ROIs with small unwanted artifacts to pass, but it is robust enough to pass a large quantity of suitable ROIs.

3.6 Measuring the NS-SFR

NS-SFRs are calculated from the selected ROIs using the ISO12233 e-SFR slanted-edge algorithm. Rather than using linear edge fitting, as described in the standard, a 3rd order polynomial fitting is used to allow curved edges to be used without introducing error [27, 28]. Burns’ *sfrmat4* MATLAB implementation [19] is used for this purpose.

Minor adjustments were made to the *sfrmat4*, which did not impact the slanted-edge algorithm itself. These allowed to output edge angle, edge contrast, and check whether clipping was detected. Adjustments were also made to flag edges that do not have a signal in all three RGB color channels. Furthermore, fault identifiers were applied in the code to flag errors in the *sfrmat4* caused by unsuitable ROIs. These ROIs are subsequently deselected.

NS-SFR results with high noise are filtered out using a 4th order polynomial fitted to each curve. If the error between the data and fit exceeds 0.1 modulation, the curve is deselected. This filtration of data further helped to remove anomalous results that potentially bias the data.

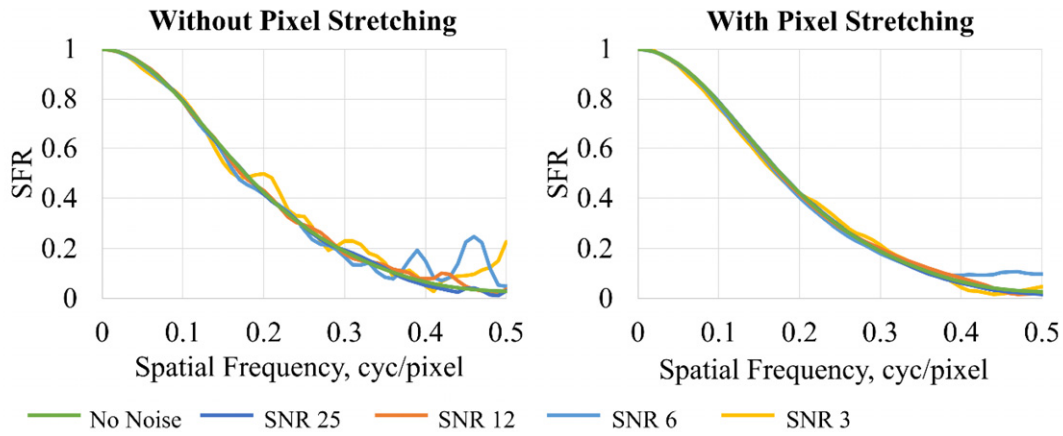


Figure 5. These two graphs illustrate the pixel stretching edge isolation’s effect on ROI with high noise, SNRs 25, 12, 6 and 3. On the left are the resultant SFRs from these ROIs without pixel stretching, and on the right are SFRs from the same ROIs but with pixel stretching.

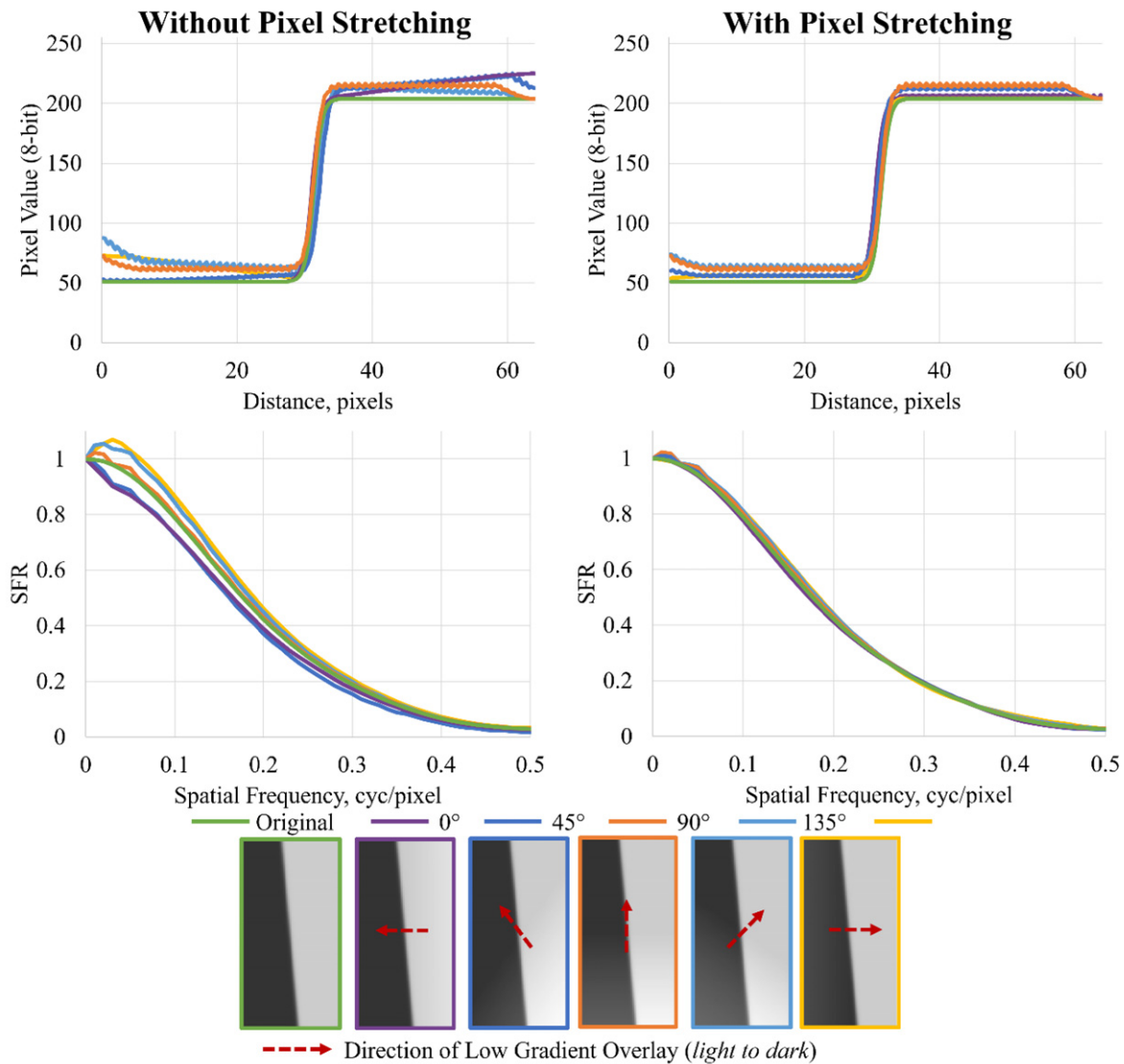


Figure 6. A demonstration of how the pixel stretching edge isolation reduces the effects of non-uniform illumination on the ISO12233 e-SFR. The left shows the ESF and SFR from ROIs with a low-gradient overlay. On the right are the same ROIs but using pixel stretching.

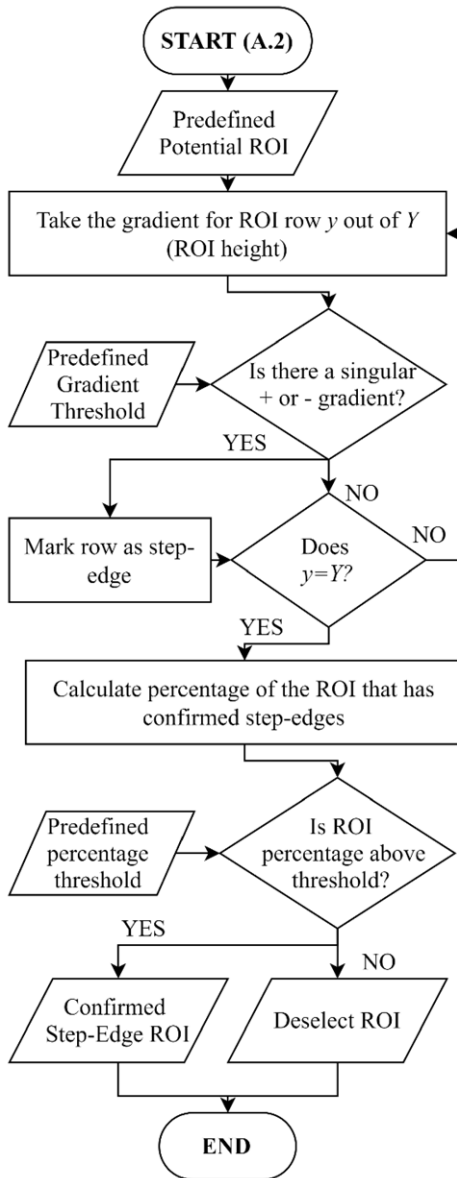


Figure 7. Step-edge ROI verification flowchart describing the process to select only step-edge ROIs. This process is marked with ‘A.2’ in the framework in Fig. 2.

The complete NS-SFR measurement process is detailed in Figure 9.

4. VARIATION IN THE SLANTED-EDGE METHOD

4.1 Variation due to Edge and ROI Parameters

NS-SFRs gathered using this framework will inherently contain data measured from a broad range of edge parameters. Since edge and ROI parameters (such as angle, contrast and size of the ROI window) are known to impact accuracy of the slanted-edge method [24, 29–33], the impact of these parameters needs to be understood.

ISO12233 recommends a 5° edge angle [1]. Previous studies showed that increasing edge angle increases error [24, 29, 30], but generally, a 3°–30° angle from the vertical is found acceptable [24].

Edge contrast was shown to have little effect on e-SFR, with no image noise present. With noise, low contrast edges become more corrupted, when compared to high contrast edges [30]. On the other hand, high contrast edges are more prone to non-linear sharpening [31, 32]. Guidelines recommend a low to medium modulation contrast of 0.55–0.65 to avoid these contradicting characteristics [1].

ROI window size has also been shown to introduce variation in e-SFR [24, 29]. The vertical ROI height determines the number of ESFs used in the edge resampling process. More rows lead to more accurate results. The minimum ROI height for a stable e-SFR depends on the slanted-edge algorithm [29]. Generally, it is recommended to be greater than 80 pixels for the highest accuracy but can be reduced to 40 pixels if there is low noise [34].

The horizontal ROI width determines the number of data points that are Fourier transformed. For low noise ROIs, the width will increase the number of e-SFR sample points without any negative consequence. A wide ROI will include more noise at higher noise levels, introducing increased high frequency positive bias to the e-SFR result. Therefore, a narrow ROI is recommended [24, 33], as long as it does not inhibit the ESF. In addition to a narrow width, a large ROI height keeps the ROI vertical-to-horizontal aspect ratio high [24], diminishing the effects of noise and other image artifacts (for example, from compression) on the resulting e-SFR.

4.2 Parameter Selection for System e-SFR Estimation

Parameter values for the system e-SFR estimation were established by assimilating knowledge from the well-documented sources in the previous sub-section and conducting e-SFR variation studies. The e-SFR parameters interaction with high noise levels and the pixel stretching filter was of particular interest.

Both simulations and captured test chart data were used to validate the edge and ROI parameter ranges. This paper report results only from simulations; the test chart data resulted in similar trends.

The ROIs were simulated with the fixed parameter baseline of 2.5°–5° edge angle, a Michelson contrast of 0.55–0.65, and an ROI height and width of 128 pixels. Edge angle, edge contrast, ROI height and ROI width ranges were expanded individually. Six noise levels were added, using a combination of Poisson and Gaussian distributions to emulate shot and read noise, respectively, resulting in SNR values 97, 76, 51, 37, 26 and 18. Including the noiseless ROIs, seven noise levels in total were examined.

Figure 10 plots the mean absolute error (MAE) introduced by changing the individual parameters at each noise level. Each ROI was simulated/isolated three times, allowing the MAE to be calculated. The MAE was measured in reference to the lowest noise ROI with a 5° edge angle, a Michelson contrast of 0.60 and a window size of 64 × 128 pixels. It was calculated from spatial frequencies 0.4–0.5 cyc/pixel, allowing the effects of high-frequency noise error with parameter variations and pixel stretching to be studied.

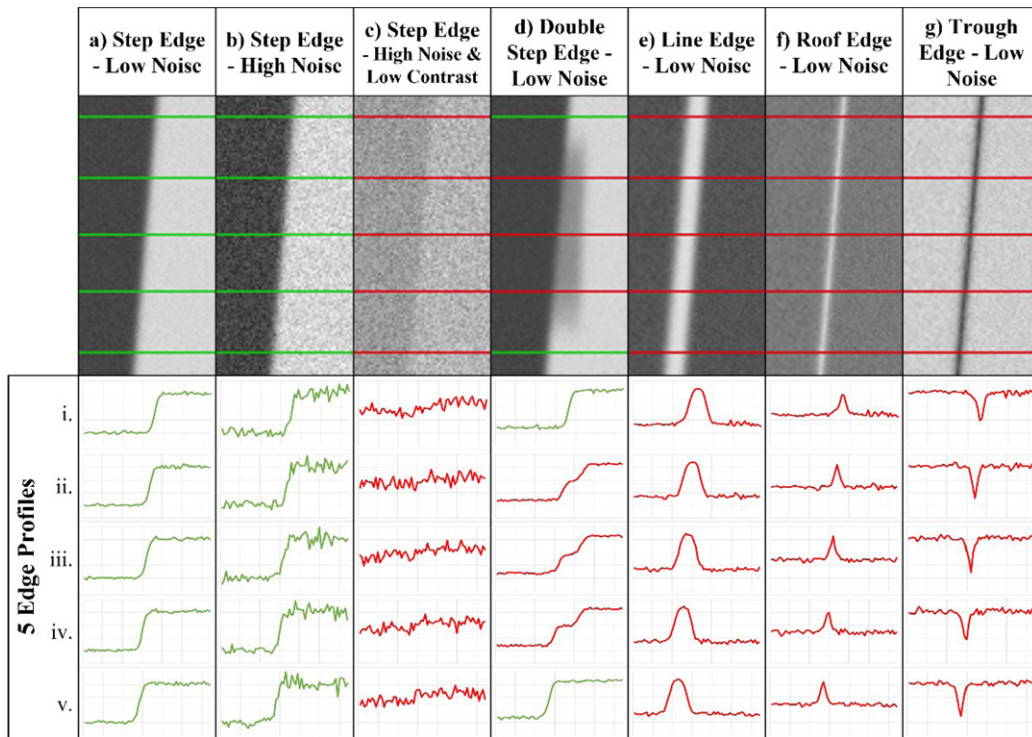


Figure 8. These seven ROIs contain various edge profiles, including step edges, staircase edge, line edge, roof edge, and a trough edge. This diagram describes how the measured gradient can successfully determine useable step edge ROIs.

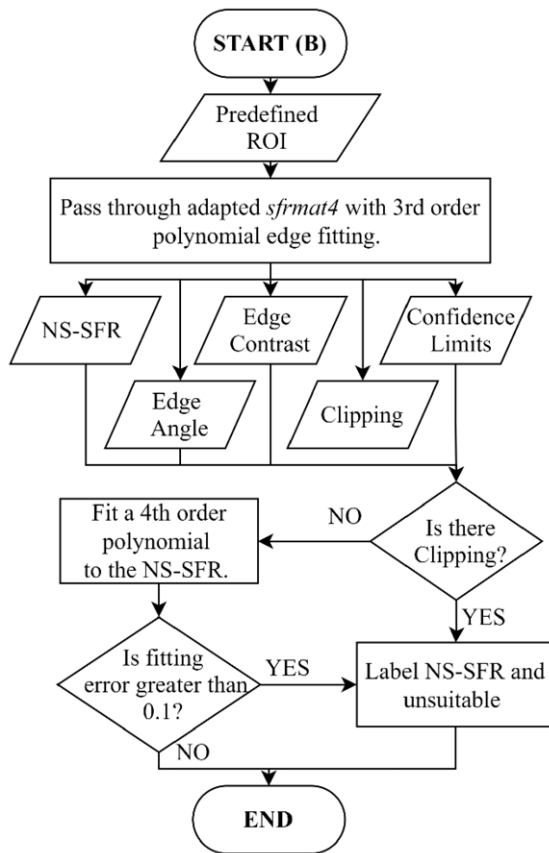


Figure 9. Flowchart depicting measuring the NS-SFR from the validated ROIs. This process is marked with 'B' in the framework in Fig. 1.

The top row illustrates variations without pixel stretching and the bottom row with pixel stretching.

This data highlights how pixel stretching reduces the effects of noise on the e-SFRs. Without pixel stretching, noise effects are apparent for low angles, low contrasts, small ROI heights, and wide ROI widths. Variation is significantly reduced with pixel stretching, providing results closer to noiseless behavior.

Pixel stretching also reduces artifacts from compression and scene textures. This 'noise reduction' allows the acceptable ranges to be maximized, selecting significantly more NS-SFR data for estimating the system e-SFR without introducing substantial MAE.

Table I lists NS-SFR parameter ranges adequate for estimating system e-SFR, along with the ISO12233 recommended parameters. Since significantly restricting the angle limits the number of edges isolated from natural scenes, edge angles were kept within a broad range (from 2.5° to 35°). Although pixel stretching improves low contrast edge MAE, it does not reduce the bias introduced by non-linear camera ISPs. Thus, contrast range was kept from 0.55 to 0.65, as per ISO12233, to restrict non-linearity effects. Variation from the ROI window size is noise dependent and is reduced by implementing pixel stretching; therefore, a broad range of ROI sizes can be implemented in the estimation. The NS-SFR framework minimum ROI crop sizes (see section 3.3) were set from the results presented here. Thus, as the framework isolates the appropriate window sizes, no NS-SFR window size restrictions are required.

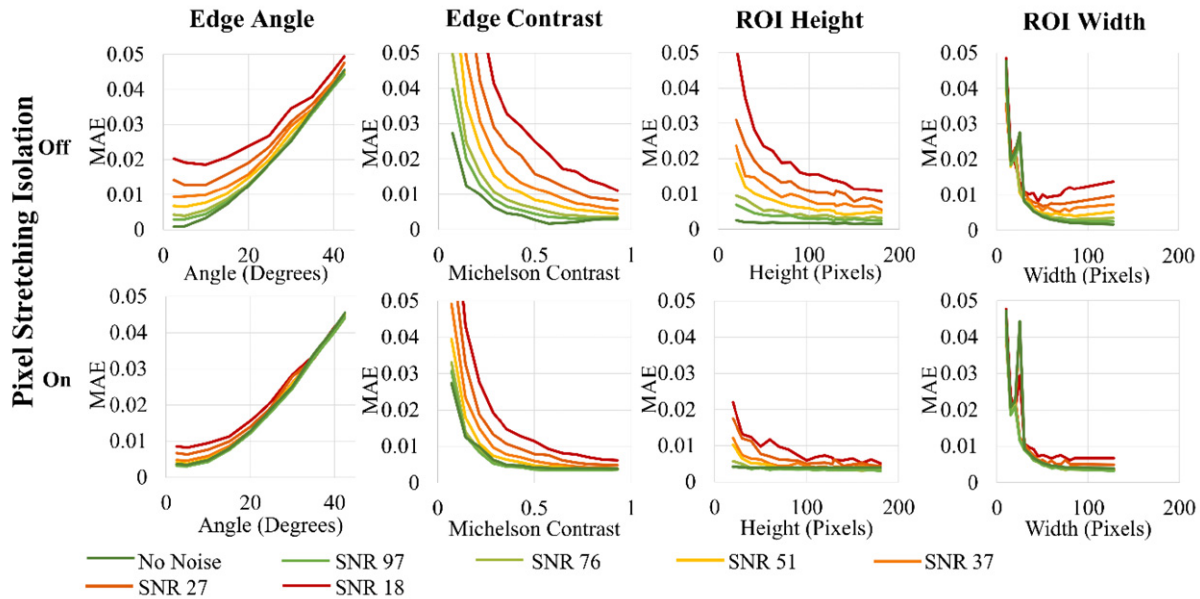


Figure 10. Comparison of the simulated MAE introduced by the SFR parameters, edge angle, edge contrast, ROI height and ROI width, with and without applying pixel stretching at seven noise levels.

5. ESTIMATION OF THE SYSTEM E-SFR

5.1 System e-SFR Estimation Methodology

The NS-SFR data derived from the image dataset is first compiled and divided into six uniform radial annulus segments (1/6 refers to the central segment 6/6 to the frame corners), visually represented in Figure 11. The number of radial annuli can be adapted depending on the application. Horizontal and vertical edge orientations are kept separate.

Next, an edge selection process is carried out to identify optimal step-edges, aiming to isolate edges with an input that maintains a higher modulation across the spatial frequencies than the camera system rendering capabilities. These edges are selected by measuring the LSF full width at half maximum (FWHM) for every ROI per radial annuli and orientation. A threshold is applied to the resultant LSF FWHM distribution to return the narrowest LSFs. The 10th percentile of the distribution is currently used for this purpose.

Fig. 11 demonstrates the relation between the NS-SFR and ISO12233 test chart LSF FWHM distributions

Table I. ROI and edge parameter ranges for the e-SFR ISO12233 method and those used to estimate the system-SFR from the NS-SFR data.

Parameter	ISO12233	NS-SFR
	e-SFR	
Edge Angle	< 45° (5° Recommended)	2.5° – 35°
Edge Contrast	0.55–0.65	0.55–0.65
ROI Size	> 64 × 80–500 pixels	> 20 × 20–128 pixels

from systems described in section 5.2. By comparing the threshold's position to the distribution from a test chart, it is seen that the 10th percentile corresponds to the narrowest ISO test chart LSF output. This threshold could be expanded to comprise more of the ISO12233 distribution; however, this would increase the probability of introducing non-optimal step-edge, incorporating the scene component of the NS-SFRs. Therefore, only expected near-optimal edges are used.

The selected NS-SFR data is arranged into a 4-dimensional binning coordinate system; each coordinate referring to a unique combination of the following parameters: spatial frequency, edge angle, edge contrast and radial annuli. Each coordinate has an associated NS-SFR value.

The binning edge parameter values are set as follows:

- angle (10 values): at intervals of 5°, ranging from 5° to 40°, plus 2.5° and 42.5°.
- contrast (21 values): at intervals of 0.05, ranging from 0 to 1 Michelson contrast.

The two NS-SFR input parameters are assigned to the closest relevant bin edge parameter value, resulting in the NS-SFR being assigned a unique set of coordinates. If there are NS-SFRs with the same coordinates, their LSFs are averaged in the spatial domain. The averaged LSF is then used to recalculate the NS-SFR for the given set of coordinates. The average was performed in the spatial domain to reduce the high frequency bias caused by image noise [35], which can compile when the SFR is averaged in the frequency domain.

Using *Piecewise Cubic Hermite Interpolating Polynomial* (PCHIP) [36, 37], the NS-SFRs are resampled to 51 samples, at spatial frequencies ranging from 0 to 0.5 cyc/pixel.

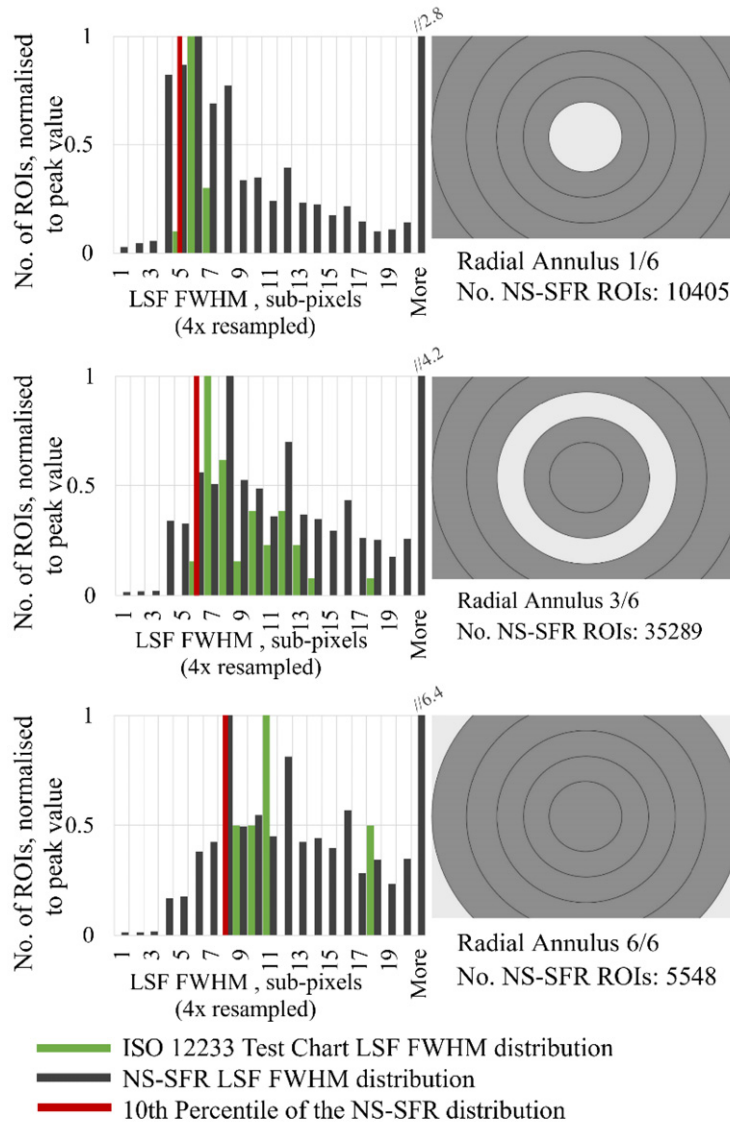


Figure 11. A comparison of the LSF FWHM distribution from NS-SFR data and ISO12233 test chart data for three radial annulus segments. The 10th percentile of the NS-SFR LSF FWHM distributions is marked. The data is from captures with the Nikon D800 DSLR and a 24 mm lens set at $f/4$.

51 frequency values provide the NS-SFRs at appropriate summary measures, such as 25%, 50% and 75% of the Nyquist frequency, as well as high resolution sample points for analysis. The number of samples can be adjusted for purpose.

NS-SFR values are binned at the given parameter coordinates and frequency range per radial annulus. The result is a 4-dimensional binning coordinate system sized $51 \times 10 \times 21 \times 6$ (spatial frequency, edge angle, edge contrast and radial annuli). Any coordinate without assigned data was set as *Not a Number* (NaN). Two coordinate systems were created, one for the vertical and one for the horizontal NS-SFRs.

The parameter ranges in Table I are used to threshold the binned NS-SFRs. This selected data is then averaged in the spatial domain per radial annulus, resulting in

six system e-SFR estimates across the frame. A weighted average of these six estimates is then calculated to obtain an overall system e-SFR estimate. Again, the average is done in the spatial domain. The applied weightages are 1.00 for the center, 0.75 for the partway regions and 0.50 for the frame's corners, corresponding to default weights in Imatest software employed for SFR analysis [38], but can be adjusted depending on the application. For example, some image quality metrics apply a heavier weight on the frame's corners than in the center, emphasizing the poorer SFRs [39].

The weighted mean is calculated from the system e-SFR estimates of the radial annuli rather than individual NS-SFRs, removing bias caused by image areas of high-density edges. As seen in Fig. 11, 3/6 (partway) region of NS-SFR data contained a greater number of isolated ROI than the center

and corners, thus taking local averages before the global weighted average removed this bias.

5.2 System e-SFR Parameter Validation

A 5th dimension was added to the system e-SFR estimation binning coordinate systems to validate the system e-SFR, the ROI height. This additional parameter was used to confirm that the ROI window size introduces negligible variation to the NS-SFR data. ROI height ranged from 30 to 130 pixels, at intervals of 10 pixels, providing 11 values.

The resulting 5-dimensional coordinates were sized $51 \times 10 \times 21 \times 11 \times 6$ (spatial frequency, edge angle, edge contrast, ROI height and radial annuli), with an NS-SFR value per coordinate.

Two sets of NS-SFR data were used. Each originated from processing image datasets from a single camera system:

- DSLR 1 system is the Nikon D800, equipped with a 24 mm lens set at $f/4$; 1866 images were captured for the dataset.
- DSLR 2 is again the Nikon D800, equipped with a 135 mm lens, also set at $f/4$; 1009 images were captured for the dataset.

Images were shot in RAW and then converted to uncompressed TIFF files, with sharpening and denoising switched off in the demosaicing process.

MAEs were calculated at each coordinate for both systems to evaluate the parameter variation in the NS-SFR data. First, the frequencies were binned, only using the high-frequencies when calculating the MAEs, ranging from 0.4 to 0.5 cyc/pixel. The MAE values were given in the form of a 4-dimensional coordinate system, sized $10 \times 21 \times 11 \times 6$ (edge angle, edge contrast, ROI height and radial annuli).

As discussed earlier (section 4), noise introduced high-frequency error. Thus, the high-frequencies illustrate the maximum MAE in the data, demonstrating the highest variation from the parameters.

The MAEs were calculated in relation to the TIFF e-SFRs derived from test chart edges under the recommended ISO12233 conditions [1]. The ISO12233 parameter ranges in Table I were set as the fixed baseline, and each parameter was then extended individually, obtaining the MAE at each interval.

Figure 12 shows MAE plotted against each of the three primary parameters for both camera systems. The results indicate that these now calibrated NS-SFRs follow expected parameter characteristics. For both systems, the error floor within the system e-SFR estimate ranges is consistent at an MAE of 0.1. The edge angle remains at the error floor up to approximately 35° . Edge contrast remains unchanging, fluctuating around the noise floor. For DSLR 1, ROI height has no variation across the entire range. DSLR 2 data has increased noise but still demonstrates a strong average nearing 0.1 MAE.

The results validate that selected parameter ranges are stable and suitable for system e-SFR estimation. Since images used for this validation originate from systems with minimal

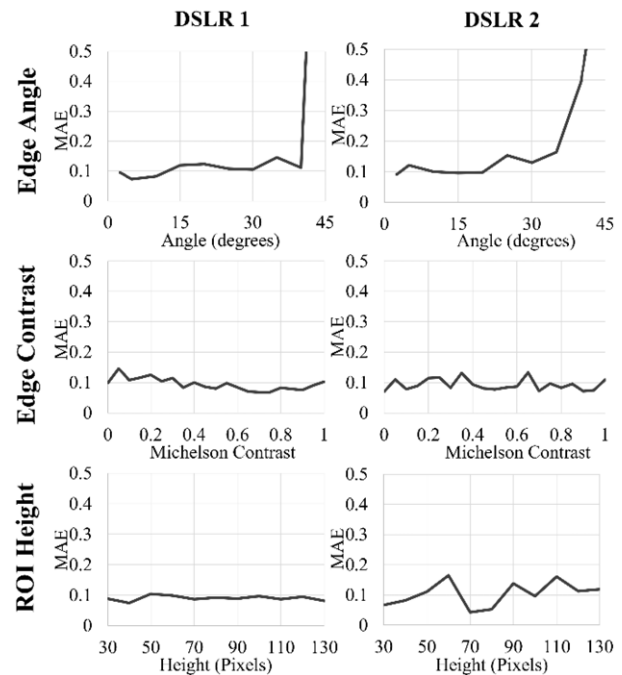


Figure 12. The comparison of the MAE introduced by the e-SFR parameters, edge angle, edge contrast and ROI height, from NS-SFRs calculated from two DSLR systems.

ISP, the plots also indicate that the edge contrast range could be increased, including more edges extracted from the natural scenes, improving system e-SFR precision. That said, the contrast range will be kept as the standard, 0.55 to 0.65 Michelson contrast, for the reasons previously stated in section 4.2.

6. SYSTEM SFR ESTIMATION RESULTS

This section presents a summary of the results and error when implementing the proposed method to derive the system e-SFR. Further qualitative analysis of the results from the implementation can be found in Ref. [4] by the same authors.

When using (near-) linear camera systems, the system e-SFR estimates are close to the ISO12233 e-SFR measurement. The same two DSLR camera systems were used here, as in section 5.2. In the DSLR 1 dataset, 135,376 ROIs were isolated, whilst from the smaller DSLR 2 dataset, 74,033 ROIs were isolated. On average, there were 73 ROIs per image measured from a total of 2875 diverse images. Note that, the pixel resolution of the sensor influences the number of ROIs isolated from the method. The sensor resolution of both datasets in this study was relatively high, at 36.3 megapixels. After applying thresholds to the ROIs, only 3.41% were used in the system e-SFR estimation calculation.

The ISO12233 e-SFR was measured for each system, obtaining target measurements against which the system e-SFR estimates are assessed. e-SFRs were measured across the frame, dividing the data into the six radial annuli and calculating the average for each segment. The weighted

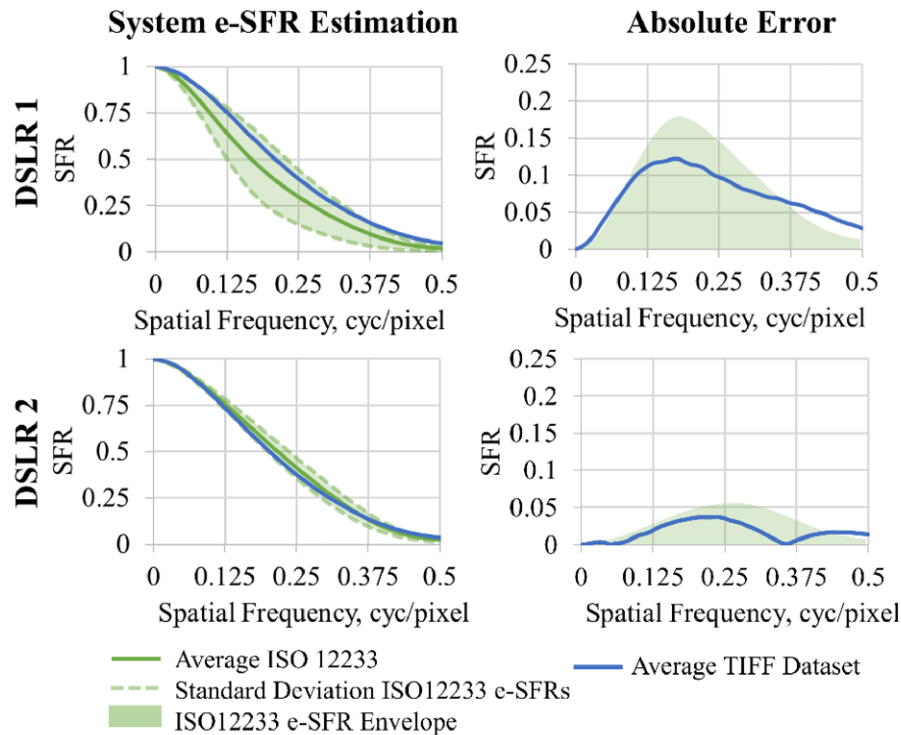


Figure 13. The weighted average system e-SFR estimations for the DSLR 1 and 2 datasets. The left column plots the estimations, and the right column the absolute error from the ISO12233 e-SFR.

average across the frame was also computed, with the same weight values as those used in the estimated system e-SFRs.

Figure 13 illustrates the weighted average estimated system e-SFRs for the two systems, plotted in relation to the equivalent ISO12233 e-SFR. This figure also plots the absolute error of these estimates in the second column. The shaded area in each graph plots the standard deviation from the ISO12233, indicating the acceptable region for the estimates, i.e., an e-SFR estimation within the limits of the equivalent test chart measure.

Excluding a bias in the high frequencies, these results are within one standard deviation of the ISO12233, indicating acceptable accuracy. The high-frequency boost further confirms that natural scene textures add noise to the NS-SFRs.

Figure 14 provides system e-SFR estimation MAE measurements for both DSLR systems at 25%, 50% and 75% of the Nyquist frequency. MAEs are given for all six radial annuli and the weighted average. Again, the standard deviations are shaded in this figure.

Some radial annuli results are overestimated, especially in system DSLR 1. The MAE, however, does not exceed 0.085 above the ground truth standard deviation. The weighted average remains consistently within the ground truth limits.

When implementing the method on a dataset captured with a highly non-linear system, the results are not comparable with an edge test chart [4, 17]. But this is expected, as the non-linear ISP sharpens isolated chart edges, whilst natural scene edges generally are not as sharpened due

to surrounding textures and denoising. These observations indicate that the NS-SFRs can potentially form a scene dependent performance measure. Splitting the dataset up per scene category indicates that the system e-SFR estimate has a scene-dependent nature [4].

7. DISCUSSION

This paper proposes the complete methodology for a natural scene derived camera system e-SFR measure that eliminates the need for test charts and laboratory conditions, with the potential to provide real-time camera performance. Relevant e-SFR parameters are investigated and verified using simulated and captured test chart edges. Indicative results from implementing this method using image datasets from two different DSLR camera systems are summarized at the end of the paper. Detailed system e-SFR results from the two DSLR camera systems and a smartphone camera system are presented in Ref. [4] by the same authors.

System e-SFRs were derived for two different DSLR camera systems. They were shown to be close approximations of the ISO12233 e-SFR measurements from test charts, captured under the recommended capturing conditions.

The pixel stretching filter employed in the framework significantly increases the quantity of useful step edges. In addition, this filter was shown to decrease the SFR bias from high-frequency noise and reduce non-uniformities. Intensive study of edge and ROI parameters also defined appropriate parameter ranges and limits that provide more

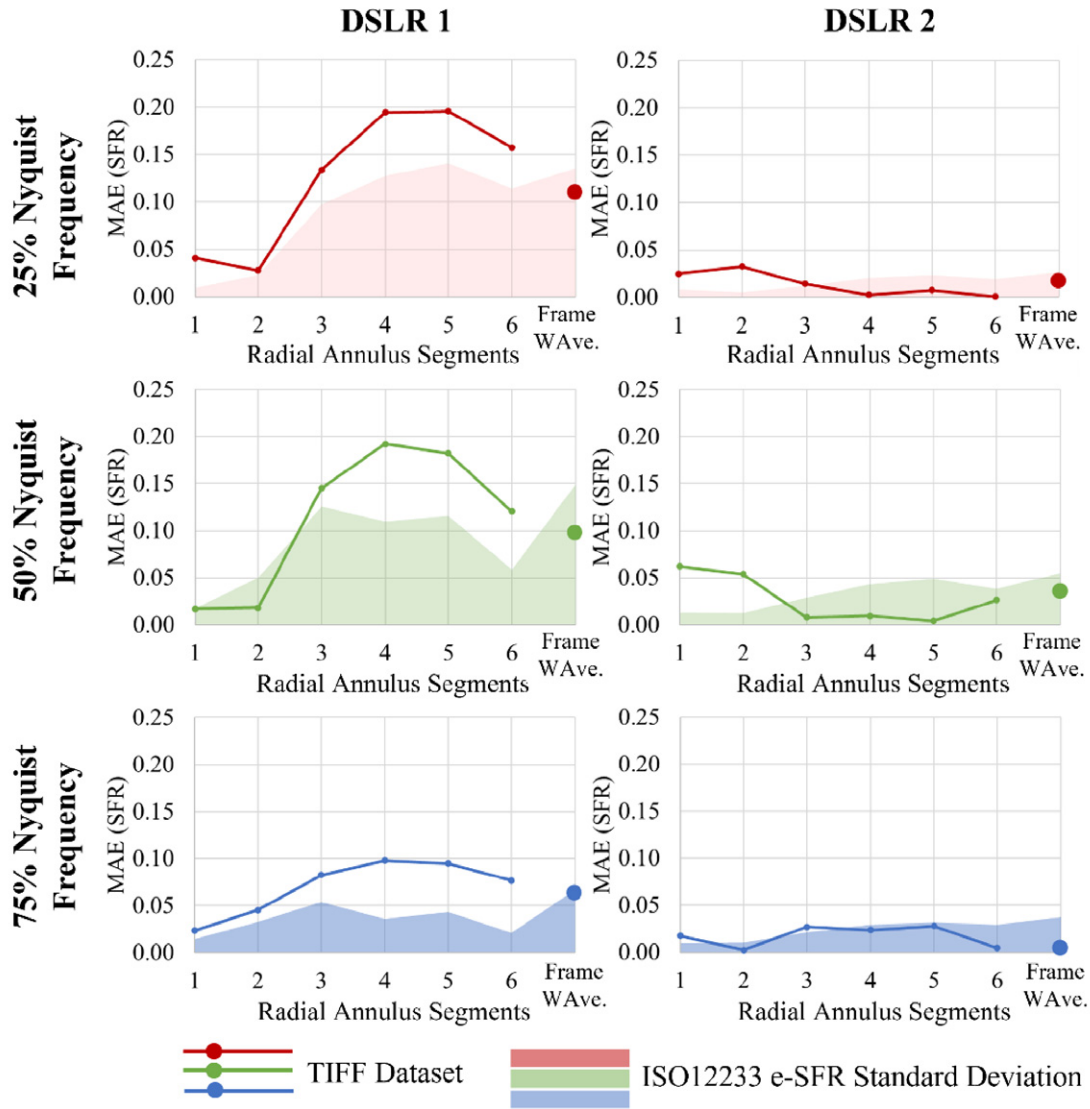


Figure 14. The MAE for two DSLR vertical system e-SFR estimates at 25%, 50% and 75% of the Nyquist frequency. This data is plotted across the six radial annulus segments and the weighted average in relation to the ISO12233 e-SFR standard deviation.

robust estimates. Consequently, the resulting estimated system e-SFR have high accuracy.

That said, several caveats and improvements need to be addressed before the proposed method is suitable for real-world applications.

The framework extracts some ROIs that contain unwanted scene information, introducing anomalous NS-SFRs. Through statistical analysis, most of these anomalies can be removed. Nonetheless, further refinements to the edge verification process will reduce these anomalies occurring. Such improvements may include replacing the static noise floor thresholds with an adaptive threshold to estimate the noise floor of individual images. This threshold can either be applied globally or at a local level, accounting for image texture in the immediate ROI area.

Although the framework isolates 73 ROI per image on average, only 3.41% are used in the system e-SFR estimation.

This low yield is due to selecting the 10th percentile of the LSF FWHM distribution and selecting only ROIs for SFR processing that matches the set edge parameter ranges. As a result, two issues arise. First, the precision decreases when fewer images (thus fewer ROIs) are used to estimate the system e-SFR [4]. Secondly, 96.59% of the ROIs isolated are not being used, yet they are still processed, increasing computational times.

The current computation time is thus extensive and impractical, averaging 20 minutes per DSLR image. It must be noted that this is based on research code as presented here, linearly passing through the two-stage process to derive system e-SFRs from natural scenes. Therefore, edge selection optimization is essential for the method to be utilized, especially if camera performance is measured in real-time. This optimization includes culling the edge selection to the

required parameters before processing, initially removing 96.59% of the data that does not need to be processed.

Training a CNN to select the appropriate step-edges would solve many current caveats. This CNN would replace step-edge detection and verification processes, selecting ROIs that will be directly used in the system e-SFR estimate. Pixel stretching will still need to be implemented. Such an approach would reduce anomalous NS-SFRs, improve ROI utilization and decrease computation times.

Although the pixel stretching filter diminishes the effects of non-uniformity on the NS-SFRs, additional non-uniformity correction can be implemented in future iterations of this method to mitigate relevant biases completely. Furthermore, as ROI uniformity is vital for an accurate e-SFR, non-uniformity correction can be found in standard e-SFR software, such as in *Imatest* [26].

The proposed method has the goal to estimate the ISO12233 e-SFR. Thus, the NS-SFRs were calculated using the standard algorithm. The only alteration outside the standard was implementing a 3rd order polynomial fitting to accommodate for curved edges. This higher-order fitting was essential to reduce bias in the NS-SFRs [27, 28], as it is rare to obtain straight edges from natural scene images.

Further improvements can be made by adjusting the standard algorithm, but at the cost of moving away from the ISO method. For example, adjusting the ESF resampling process to better conform to the projection down the edge slope was shown to improve the e-SFR measure in multiple studies [29, 30, 40, 41]. These studies provided results with reduced bias introduced by image noise, edge angle variations and ROI size variations. The approaches of these papers differ, with varied impact on comparability and compatibility with the standard. *Imatest* proposes an LSF angle correction, which is not intrusive on the standard [30]. In contrast, the Reverse Projection method remodels the ESF resampling process, giving more stability to the results across large edge parameter ranges than other slanted edge method algorithms [29]. However, the resultant Reverse Projection e-SFR are consistently underestimated in comparison. Implementing such alternatives would significantly improve the NS-SFR stability, allowing more ROIs to be used to estimate the system e-SFR.

The proposed method is dependant on the input natural scene edge; large quantities of blurred edges would change the shape of the LSF FWHM distribution, thus selecting wider LSFs in the 10th percentile threshold. Further work should study the statistical analysis of the NS-SFR extrema to estimate the system e-SFR.

Once the caveats of this method have been fully addressed, there is potential utilization in several areas of imaging science.

The simplest application would be camera system performance comparisons. A system e-SFR estimate from natural scenes would allow multiple systems to be assessed by collecting a dataset of similar images from each system. The camera itself would not necessarily need to be at hand; the images could be collected through online sources and social

media, categorizing them via the available metadata. Many of these sources are often resized, compressed, and stripped of some metadata; therefore, they should be carefully selected or considered in the application. If large quantities of images can be acquired, the result is more representative of the ISO12233 e-SFR. In addition, the datasets could be sub-classed to determine the performance under various illuminations, scenes, and settings. Only minor modifications to the method would be required.

As mentioned previously, the system e-SFR estimation has shown scene dependent characteristics. With further analysis to study the scope of this scene dependency, the NS-SFRs may be used in a similar fashion to the SPD-MTF, as inputs to spatial IQMs [10–12]. This measure is based on the slanted-edge method; thus, it will be most sensitive to non-linear sharpening, unlike scene texture-MTFs which can effectively quantify denoising. Further work also is needed to evaluate the suitability of a measure based on NS-SFRs as inputs to IQMs, perhaps compared to the CNN PSF estimation method [14]. Such a study would research the NS-SFRs based on different contrast and LSF FWHM threshold ranges.

The system e-SFR estimate could potentially evaluate camera performance in real-time, giving a live-SFR feed. Such a measure can be utilized in autonomous vision systems, as discussed in Ref. [17] by the same authors. In addition to refining some processes, the method should also be optimized with relevant datasets.

8. CONCLUSION

In this paper, a framework to measure Natural Scene derived SFRs (NS-SFRs) is detailed. Further, a method for estimating the standard ISO12233 e-SFR, the system e-SFR from the resulting NS-SFRs, is also described.

NS-SFRs are measured using a series of logical procedures to detect, isolate and verify step-edges from natural scene images. First, edge detection is used to locate all image edges from an image. Next, appropriately sized ROIs are cropped, and the edges of interest isolated from other unwanted scene content, using pixel stretching. In addition to improving edge isolation, the pixel stretching filter reduces the effects of noise and non-uniformities. Each ROI undergoes a verification process, taking the gradient across the edge. Selected ROIs are then passed through the standard slanted-edge method algorithm to obtain a series of NS-SFRs for each image.

The system ISO12233 e-SFR is subsequently estimated using the NS-SFRs of large image datasets. Due to optical performance variations across the field of view, NS-SFR data is segmented into six radial annuli, allowing for six e-SFR estimates across the frame. Since a step-edge input containing higher modulation than the system rendering capabilities is required for the e-SFR estimation, the highest performance edges are selected for the calculation. This selection is achieved by thresholding the top 10th percentile of the LSF FWHM distribution, i.e., the narrowest LSFs. The threshold data is filtered to obtain the ROIs with edge

parameter ranges shown to introduce the least variation to the ISO12233 e-SFR. As a result, only 3.41% of the NS-SFRs are selected when implementing the proposed method. When averaged, they are stable, providing an estimate of the system e-SFR across all the frame segments comparable to that provided by the ISO standard using test chart edges in laboratory settings. When a weighted average is taken across the entire frame, the result is an accurate estimate within one standard deviation of the ISO12233.

MATLAB CODE

The latest version of the MATLAB code can be found at: <https://github.com/OliverVZ11/NS-SFR>

ACKNOWLEDGMENT

We wish to show our appreciation to Dr. Jae Young Park for funding this research project.

REFERENCES

- British Standard Institute, "BS ISO 12233:2017 Photography - Electronic still picture imaging - Resolution and spatial frequency responses," (BSI Standards Publication, 2017), pp. 1–62.
- R. Jenkin, "Noise, sharpness, resolution and information," in *The Manual of Photography*, edited by E. Allen and S. Triantaphillidou, 10th ed. (Focal Press, Taylor & Francis Group, New York & London, 2009), pp. 433–456.
- J. C. Dainty and R. Shaw, "Fourier transforms," *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes* (Academic Press, New York, 1974), pp. 190–231.
- O. van Zwanenberg, S. Triantaphillidou, A. Psarrou, and R. Jenkin, "Analysis of natural scene derived spatial frequency responses for estimating the camera ISO12233 slanted-edge performance," *J. Imaging Sci. Technol.* **65**, 060405-1–060405-12 (2021).
- B. Jayachandriah, S. Jayabharathi, S. Muralikrishnan, and A. S. Kumar, "Development of advanced data quality evaluation tools for high resolution IRS-Satellite data with IDL/ENVI," *ENVI/IDL User Mfg.* (2005), pp. 1–8.
- L. He, P. Michel, S. Puglia, and D. Williams, "Computer assisted image analysis for objective determination of scanning resolution for photographic collections – An automated approach," *IS&T Archiving Conf. Proc.* (IS&T, Springfield, VA, 2013), pp. 43–47.
- P. D. Burns, "Refined measurement of digital image texture loss," *Proc. SPIE* **8653**, 86530H (2013).
- J. McElvain, S. P. Campbell, J. Miller, and E. W. Jin, "Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems," *Proc. SPIE* **7537**, 75370D (2010).
- R. Branca, S. Triantaphillidou, and P. Burns, "Texture MTF from images of natural scenes," *IS&T Electronic Imaging: Image Quality and System Performance XIV* (IS&T, Springfield, VA, 2017), pp. 113–120.
- E. W. S. Fry, S. Triantaphillidou, R. E. Jacobson, J. R. Jarvis, and R. B. Jenkin, "Bridging the gap between imaging performance and image quality measures," *IS&T Electronic Imaging: Image Quality System Performance XV* (IS&T, Springfield, VA, 2018), pp. 231-1–231-6.
- E. W. S. Fry, S. Triantaphillidou, R. B. Jenkin, J. R. Jarvis, and R. E. Jacobson, "Validation of modulation transfer functions and noise power spectra from natural scenes," *J. Imaging Sci. Technol.* **9**, 60406-1 (2019).
- E. W. S. Fry, S. Triantaphillidou, R. B. Jenkin, R. E. Jacobson, and J. R. Jarvis, "Scene-and-process-dependent spatial image quality metrics," *IS&T Electronic Imaging: Image Quality and System Performance XVII* (IS&T, Springfield, VA, 2019), pp. 60407-1–60407-13.
- Imatest, "Spilled-Coins Dead Leaves, and Random Chart Analysis Scale-invariant test charts (v.2021.1)," *Imatest*, 2021 <http://www.imatest.com/docs/random/> (accessed Jul. 05, 2021).
- M. Bauer, V. Volchkov, M. Hirsch, and B. Schöckkopf, "Automatic estimation of modulation transfer functions," *IEEE Int'l. Conf. on Computational Photography (ICCP)* (IEEE, Piscataway, NJ, 2018), pp. 1–40.
- O. Van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Edge detection techniques for quantifying spatial imaging system performance and image quality," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops* (IEEE, Piscataway, NJ, 2019), pp. 1871–1879.
- O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Camera system performance derived from natural scenes," *IS&T Electronic Imaging: Image Quality and System Performance XVII* (IS&T, Springfield, VA, 2020), pp. 241-1–241-10.
- O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Natural scene derived camera edge spatial frequency response for autonomous vision systems," *IS&T London Imaging Meeting* (IS&T, Springfield, VA, 2021).
- J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986).
- P. D. Burns, "sfrmat4 source code," (Burns Digital Imaging, 2020) (accessed Oct. 04, 2020).
- R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision* (Addison-Wesley, Boston, MA, 1992), Vol. II, pp. 316–317.
- J. P. Lewis, "Fast normalized cross-correlation," *Vision Interface* **1995**, 1–7 (1995).
- Y. Huo, G. Wei, Y.-D. Zhang, and L. Wu, "An adaptive threshold for the canny operator of edge detection," *2010 Int'l. Conf. on Image Analysis and Signal Processing* 2010, pp. 371–374.
- MATLAB, "regionprops" *MathWorks*, 2006 <https://uk.mathworks.com/help/images/ref/regionprops.html> (accessed Jul. 28, 2021).
- D. Williams, "Benchmarking of the ISO 12233 Slanted-edge Spatial Frequency Response Plug-in," *IS&T PICS 1998: Image Processing, Image Quality, Image Capture Systems Conf.* (IS&T, Springfield, VA, 1998), pp. 133–136.
- D. Williams and P. D. Burns, "Evolution of Slanted Edge Gradient SFR Measurement," *Proc. SPIE* **9016**, 901605 (2014).
- Imatest, "Correcting non-uniformity in slanted-edge MTF measurements," *Imatest*, 2020 (accessed Mar. 10, 2021).
- P. D. Burns and D. Williams, "Camera resolution and distortion: Advanced edge fitting," *IS&T Electronic Imaging: Image Quality and System Performance XV*, pp. 171-1–171-5 (IS&T, Springfield, VA, 2018).
- P. D. Burns, D. Williams, J. Griffith, H. Hall, and S. Cahall, "Application of ISO standard methods to optical design for image capture," *IS&T, Electronic Imaging 2020 Image Quality and System Performance* (IS&T, Springfield, VA, 2020), pp. 240-1–240-7.
- S. Birchfield, "Reverse-Projection Method for Measuring Camera MTF," *IS&T Electronic Imaging: Image Quality and System Performance XIV* (IS&T, Springfield, VA, 2017), pp. 105–112.
- J. K. M. Roland, "A study of slanted-edge MTF stability and repeatability," *Proc. SPIE* **9396**, 93960L (2015).
- U. Artmann, "Image quality assessment using the dead leaves target: experience with the latest approach and further investigations," *Proc. SPIE* **9404**, 94040J (2015).
- Y. Okano, "Influence of non-linear image processing on spatial frequency response of digital still cameras," *Opt. Rev.* **5**, 358–362 (1998).
- P. D. Burns and D. Williams, "Refined slanted-edge measurement for practical camera and scanner testing," *IS&T PICS 2002: Image Processing, Image Quality, Image Capture Systems Conf.* (IS&T, Springfield, VA, 2002), pp. 191–195.
- Imatest, "Using SFR part 2: Running Imatest SFR." (accessed Dec. 04, 2020).

- ³⁵ E. C. Yeadon, R. A. Jones, and J. T. Kelly, "Confidence limits for individual modulation transfer function measurements based on the phase transfer function," *Photogr. Sci. Eng.* **14**, 153–156 (1970).
- ³⁶ F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM J. Numer. Anal.* **17**, 238–246 (1980).
- ³⁷ D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software* (Prentice Hall, Upper Saddle River, NJ, 1988).
- ³⁸ Imatest, "[Using eSFR ISO Part 2.](#)" (accessed Dec. 07, 2020).
- ³⁹ BSI Publication, "BS ISO 20462-3:2012 Psychophysical experimental methods for estimating image quality Part 3: Quality ruler method," *BSI Standards Publication* (2012), pp. 1–23.
- ⁴⁰ K. Masaoka, T. Yamashita, Y. Nishida, and M. Sugawara, "Modified slanted-edge method and multidirectional modulation transfer function estimation," *Opt. Express* **22**, 6040 (2014).
- ⁴¹ H. Li, C. Yan, and J. Shao, "Measurement of the modulation transfer function of infrared imaging system by modified slant edge method," *J. Opt. Soc. Korea* **20**, 381–388 (2016).