Angle- and Noise-Robust Spatial Frequency Response-based Resolution Analysis in Natural Scenes

Sara Lee[†], Seungwan Jeon[†], Subin Han, Yu Gyeong Lee, KiChul Park, Sung-Su Kim; S.LSI Division, Samsung Electronics; Hwaseong-si, Gyeonggi-do, Republic of Korea

[†]these authors contributed equally to this work

Abstract

Evaluating spatial frequency response (SFR) in natural scenes is crucial for understanding camera system performance and its implications for image quality in various applications, including machine learning and automated recognition. Natural Scene derived Spatial Frequency Response (NS-SFR) represented a significant advancement by allowing for direct assessment of camera performance without the need for charts, which have been traditionally limited. However, the existing NS-SFR methods still face limitations related to restricted angular coverage and susceptibility to noise, undermining measurement accuracy. In this paper, we propose a novel methodology that can enhance the NS-SFR by employing an adaptive oversampling rate (OSR) and phase shift (PS) to broaden angular coverage and by applying a newly developed adaptive window technique that effectively reduces the impact of noise, leading to more reliable results. Furthermore, by simulation and comparison with theoretical modulation transfer function (MTF) values, as well as in natural scenes, the proposed method demonstrated that our approach successfully addresses the challenges of the existing methods, offering a more accurate representation of camera performance in natural scenes.

Introduction

Evaluating a spatial frequency response (SFR) in images is crucial for understanding camera system performance by quantifying how well a camera depicts the details and its implications for image quality in various applications, including machine learning and automated recognition.

Traditionally, objective measurements of SFR also have been conducted using the edge-based spatial frequency response (e-SFR) algorithm [1], a standardized method based on the ISO 12233 which has been constantly evolving up to ISO 12233:2024 [2] now, utilizing slanted edges input for evaluation. The e-SFR measurement involves identifying regions of interest near slanted edges, computing the edge spread function (ESF) by projecting pixels along specific direction such as the direction perpendicular to the edge, deriving the line spread function (LSF) through differentiation, and finally applying a Fourier transform to the LSF to obtain the SFR or the modulation transfer function (MTF) [1]. The previously mentioned assessments of e-SFR basically relied on chart-based method, literally requiring the use of predetermined slanted edge patterns for evaluation, thereby restricting its applicability to common scene images, which often involve slanted edges.

Fortunately, the introduction of Natural Scene Spatial Frequency Response (NS-SFR) [3] has made it possible to evaluate the camera performance in natural environments without the need for specialized charts. With the emergence of this approach, it has become possible to move beyond fixed chart-based analysis and extend camera performance assessment to general natural scene. While it fundamentally relies on the previous ISO 12233 e-SFR, it introduced some additional preliminary steps before deriving e-SFR by focusing on detecting and extracting suitable step edges from natural scene images [3].

Whereby the NS-SFR proposed by Van Zwanenberg [3] divided the process into five main stages, it also may be broadly categorized into the following three functional aspects. The first is edge extraction, the second is region of interest (ROI) isolation & validation, and the third is MTF calculation. The first stage includes the use of the Canny algorithm to detect edges from the entire image [3]. The second stage involves extracting candidate ROI regions centered around the edges, performing edge isolation techniques and image processing such as pixel stretching, and validating the edges based on criteria such as angle, contrast, ROI size, and linearity and whether they are step edges [3]. The final stage is based on the traditional slanted-edge algorithm, the line spread functions (LSFs) from the edge ROIs are averaged into a single LSF considering several factors (such as the radial distances and the rankings of the MTF), and then the its MTF, so called mean MTF, is calculated in the end to represent the overall image resolution [3].

Although many studies on the NS-SFR and its associated parameter [1, 3, 4] have been conducted, the existing NS-SFR have consistently showed limited performance in special conditions especially according to the angle and noise in the extracted ROIs. To ensure reliable results, some existing methodologies [1, 3, 5] impose restrictions on the angles of edges or recommend an acceptable range for it, by excluding ROIs with angles greater than specific degrees such as 35° due to their negative impact on accuracy, limiting its potential for comprehensive analysis. Given the characteristics of natural scenes, additionally, it is common for noisy images to be inputted, resulting in significant errors in the NS-SFR results which is vulnerable to noise, as noise can affect edge detection and the accuracy of MTF measurements.

In order to overcome the previously mentioned limitations, this paper propose a method for broader angle coverage and enhanced noise robustness of NS-SFR. Regarding the limited angle range, an adaptive oversampling rate (OSR) [6] and phase shift (PS) [7] method (hereafter referred to as an adaptive binning) suggested by Wu et al. [8] has been developed for accurate assessment of slanted edges at angles beyond 35 degrees. By applying this edge method in existing NS-SFR [3], we improved the MTF assessment across a broader range of angles. Even with the adaptive OSR and PS method, however, it is still insufficient to resolve the error issues caused by noise. Thus, we further implemented the Adaptive Windowing technique to improve noise robustness in images.

The overall structure of this paper is as follows. First, based on the background knowledge of the e-SFR and NS-SFR discussed earlier [1, 3] we provide detailed methods of the Adaptive Binning and Adaptive Windowing techniques to enhance reliability of the MTF measurement in NS-SFR. Next, to evaluate the effectiveness of the proposed approach, simulations comparing individual MTF analysis results against theoretical values and also verification in real images is conducted. Finally, we conclude by discussing implications of our method and suggesting future work directions to further enhance NS-SFR area.

Methods

Based on the existing NS-SFR [3] framework mentioned previously, our proposed method introduces key modifications to enhance result reliability while minimizing dependence on angle and noise levels. In the following sections, we will discuss the specific details regarding angle and noise level aspects, as well as other additional methods to improve reliability of MTF measurement in natural scene.

Extend Angle Coverage

The edge profile, which is also referred to as ESF (Edge Spread Function) is reconstructed by projecting pixel values into 1D supersampled array [1]. During the construction of the edge profile, it is crucial to select appropriate values for the oversampling rate(OSR) and phase shift (PS) in the binning method. The optimal adaptive binning method proposed by Wu et al. [8] offers a solution to overcome the limited range of edge angles. By implementing this approach into the NS-SFR, we were able to extend the angle coverage and enhance the accuracy of the results.

According to this study [8], standard edge method in ISO 12233: 2017 [1], with an oversampling rate fixed at 4, offers the benefit of low computational cost while ensuring valid data within the Nyquist limit. However, it becomes unsuitable for extremely small or large edge angles. In this perspective, it applies an adaptive binning method when extracting edge spread function (ESF), dynamically adjusting the binning width and its number of PS according to each edge angle [8]. To enhance accuracy, multi-phase binning was also employed as an additional method [8].

The specific OSR, PS values are determined by the following Equation (1).

$$(\text{OSR}, N_{\text{PS}}) = \left\{ \begin{array}{ll} (8,8), & \text{if } \theta \in \left[0, \arctan \frac{1}{18}\right) \text{deg} \\ \left(\frac{1}{2\tan(\theta)}, 6\right), & \text{if } \theta \in \left[\arctan \frac{1}{18}, \arctan \frac{1}{9}\right) \text{deg} \\ \left(\frac{1}{\tan(\theta)}, 4\right), & \text{if } \theta \in \left[\arctan \frac{1}{9}, \arctan \frac{1}{4}\right) \text{deg} \\ \left(\max \left[\frac{1}{\tan(\theta)}, 2\right], 6\right), & \text{if } \theta \in \left[\arctan \frac{1}{4}, \arctan 1\right) \text{deg} \end{array} \right.$$

(1)

By applying this adaptive value set in the process of NS-SFR, we were able to effectively gather sufficient sampling data while reducing computational cost, even for edges with extreme angle which are highly likely to exist in natural scene.

In addition, in order to make the NS-SFR more suitable for real-world images, i.e., to prevent the distortion in the edges from being more significantly reflected as the size of the ROI increases, in this study we applied a reduction from the existing maximum ROI height of 128 to 64. The validity of this value will be discussed in detail in the results section.

Improve Noise Robustness

Adaptive Windowing

To minimize the side effect, we applied an adaptive window (see Figure 1) based on Tukey with different widths depending on the edge width to tightly envelope the LSF by the use of Equation (2).

$$w_{Tukey}(x,h) = \begin{cases} 1, & \text{if } |x| < \tau_1 \\ 0.5 + 0.5 \cos\left(\frac{|x| - \tau_1}{|\tau_2 - \tau_1|}\right), & \text{if } \tau_1 \le |x| < \tau_2, \\ 0, & \text{elsewhere} \end{cases}$$

where $\tau_1 = \text{edge width} = 4s$
 $\tau_2 = 2\tau_1$ (2)

In particular, we extracted edge width information from ESF and applied an adaptive Tukey window instead of the fixed Tukey window which is applied to LSF when extracting SFR according to the flowchart depicted in Figure 1.



Figure 1. Flowchart illustrating the process of applying an adaptive Tukey window during the extraction of the SFR from edge in ROI: the window equals for 1 for $|x| < \tau_1$ and tapers to 0 for $|x| > \tau_2$, factor τ_1 and τ_2 which are specifically described in Equation 2.

Additionally, prior to calculating the edge width, noise was reduced by using Non-local means (NLM) denoising to the ROI, specifically to improve the accuracy of edge width measurement. Note that we inferred an approximate noise sigma, σ_n , of each ROI using "estimate_sigma" function in skimage python library, and empirically determined the NLM strength to be set at 4n.

The ESF from the denoised ROI was then fitted to an error function *s* in Equation (3) using logistic regression. In Equation (3), v(x) is ESF with an edge located at x = 0, and (v_{left}, v_{right}) represent the brightness levels on each left and right side of the edge in the ESF. Additionally, the optimized z-axis scale (s) is related to the edge width, with the relationship of the edge half-width, $h \approx 1/s$.

$$s = \underset{s,v_{right},v_{left}}{argmin} \left| \left(\frac{v_{right} - v_{left}}{2} erf(sx) + \frac{v_{left} + v_{right}}{2} \right) - v(x) \right|_{2}$$
where $erf(t) = \frac{2}{\sqrt{\pi}} \int_{0}^{t} e^{-z^{2}} dz$ (3)



Figure 2. The ESF graph fitted with an error function s. The width of the ESF can be used to define a new z-axis scale (s) for resampling the ESF.

Based on the edge width derived from this error function as shown in Figure 2, a window with a cosine tapered region, as previously mentioned, determined by a multiple of the measured edge width is finally applied.

Additional Approaches

In addition to the aforementioned methods, we also applied several supplementary approaches to improve noise robustness during the ROI selection.

Specifically, a simple denoising step composed of medianBlur and bilateralFilter in OpenCV python library was added as a preprocessing before running the Canny detector, which is especially used to identify the edge locations from the natural scene. By applying this complementary step only used for detection, we aimed to extract necessary candidate edges, even in highly noisy conditions.

Furthermore, after the conventional pixel stretching [3] which is conducted to exclusively isolate edge regions, we additionally added an additional brightness rescaling step. In brightness rescaling, we rescaled the pixels in a row to set both ends into specific fixed values which are same in all rows. These fixed values were set to the mean pixel values at both ends across all rows as depicted in Figure (3).



Figure 3. Illustration of the effect of applying brightness rescaling: before and after.

While the pixel stretching [3] removes nearby edges or artifacts by stretching the pixels horizontally, incorporating this method helped reduce distortion in the vertical direction, particularly being effective in reducing stripe patterns or the impact of unwanted highfrequency signals that occurred after pixel stretching in the presence of significant noise or brightness gradients.

Moreover, since the edge is analyzed with the height differences of an edge, we considered the noise level criteria redefined with Contrast-to-Noise Ratio (CNR) in Equation (4), where S_{Left} , S_{Right} are signal intensity for each side of edge and σ_{Noise} being standard deviation of image noise rather than SNR to be more suitable. In practice, we used this CNR when validating edge in perspective of noise.

$$CNR = \frac{\left|S_{Left} - S_{Right}\right|}{\sigma_{Noise}} \tag{4}$$

Additional Improvements

Continuous Weighting Factor

Before extracting the average MTF which represents the whole image resolution, LSFs from the validated edge ROIs are averaged into a single LSF with corresponding weight factor. To ensure that the weight factor of each ROI accurately reflects its feature and the average MTF value correctly represents the actual image information, we decomposed the weight factor into more detailed components, focusing on three factors in this paper.

The first factor is determined by field distance. In the previous method [3, 9], weight value according to field distance of each ROI was applied such that the weight decreases with increasing distance from the center in a discrete manner, whereas in our approach, we applied weight factor continuously. The second factor is determined by edge width, using the z-score of the distribution induced by edge width to assign a greater weight as it narrows. And the third factor is determined by z-scores of the distribution induced by deviation between each ESF and the pseudo-average ESF, primarily calculated by using the first and second factors computed earlier. Finally, the three aforementioned factors are multiplied together and applied to each ROI to derive final average LSF.

The methods used in conventional and proposed are summarized in the following Table 1 and Table 2, respectively:

 Table 1. Comparison of Conventional and Proposed methods

 for ROI Extraction before MTF Calculation

	Conventional	Proposed	
ROI Max Height	128	64	
Noise Level	SNR 5	CNR 5	
Angle	[2.5, 35]	[2, 44]	
Brightness Rescaling	OFF	ON	

Table 2. Comparison o	f Conventional and	Proposed methods
for MTF Calculation		

	Conventional	Proposed	
Over Sampling Rate (OSR) in Projection	Fixed OSR(4)	Adaptive OSR(2~8)	
Phase in Projection	1 Phase	6~8 Phase	
Window	Tukey Window	Adaptive Window	
Averaging	Only certain percent(%) of narrowest LSFs counted	Every LSFs counted with Proposed Weighting method	

Measurement of Accuracy

Simulation

To evaluate the measurement accuracy of the MTF, we primarily assumed that an ideal step edge could be represented as the Heaviside step function. Subsequently, as it passes through the blurring system such as camera imaging system, we anticipated that the ideal MTF approximates an LSF with a 1D Gaussian profile as Equation (5).

$$MTF_{ideal}(f) = e^{-2\pi^2 \sigma_b^2 f^2}, \sigma_b : blur \ sigma \tag{5}$$

We conducted simulations under various conditions by controlling edge angle, contrast, blur, and noise parameters. Specifically, the sharpness was adjusted using Gaussian blurring with the blurring sigma, σ_b . Based on the assumption of the ideal MTF, we quantified the MTF accuracy by calculating the root mean square error (RMSE) between the simulated MTFs calculated by applying the conventional or proposed methods and the ideal MTFs as in Equation (6) where k refers to spatial frequency, and N indiciating the number of all components along the spatial frequency of MTF.

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (MTF_{simul}(k) - MTF_{ideal}(k))^2}$$
(6)

Real scene images

We also evaluated the performance in real scene, to verify the MTF reliability at various noise levels that can realistically occur in natural enivornments by varying effective integrated time (EIT), which details will be discussed later in the result section. In particular, in images with strong sharpening, the SFR profile can be significantly affected by undershooting and overshooting, and in such cases, the masking width of our adaptive window is applied sub-optimally, we used input images without any additional ISP processing other than simple demosaicing from raw data.

Results

Simulation: Edge Angle

First, we analyzed the results of simulation in terms of edge angles. We assessed the RMSE between the ideal MTFs of the conventional and proposed methods across various sharpness and angle (Figure 4). For angles between 2° and 44° , the average RMSE values decreased by more than half, with values from 0.0201 to 0.0078 in noise free condition (Figure 4a) and with values from 0.0627 to 0.0275 even in the presence of a certain level of noise with CNR below 50 (Figure 4b), respectively, indicating that the proposed method yielded accurate results for overall angles, except for the vertical and horizontal orientations. Upon closer examination of the MTF from a specific condition as shown in the last column of Figure 4, the graph from the proposed method closely resembles the ideal one. In this case, the size of ROI was 25×25 pixels and brightness on either side, divided by edge, were 150 and 100, respectively, based on an 8-bit image.

Simulation: Noise in ROI

Also we conducted a simulation in terms of noise in ROI. As mentioned earlier, as the angle approaches 45 degrees, with the conventional Windowing, the edge becomes less tightly cropped, making the edge profile calculation more vulnerable to noise. So to begin with, we examined the LSF and MTF calculation of an image closer to 45 degrees in greater detail under two conditions: applying the conventional fixed Tukey Window and applying our adaptive window to same ROI with noise (Figure 5).

In this Figure 5, the conventional method failed to extract a valid MTF graph due to the presence of noise on both sides of the edge signal in the LSF graph. In contrast, the proposed method effectively removes the noise in the LSF, enabling the generation of a valid MTF graph from close to 45-degree edge image.

Moreover, by varying the CNR value which is noise level in the simulation, it was found that the proposed method (orange graph, Figure 6a) is less vulnerable to noise than the conventional method (blue graph, Figure 6a). As shown in Figure 6b, the error RMSEs in the proposed method was reduced by approximately one-third to one-half compared to the conventional approach, across various range of noise level, indicating enhanced noise robustness of NS-SFR.



Figure 6. (a) ROIs and their corresponding MTF graphs measured by the conventional and proposed method according to CNR. (b) Comparison of the RMSEs between the measured and ideal MTF (AVG±STD, n=10) according to CNR. MTF, modulation transfer function; and RMSE, root mean square error.

In addition, by examining the results of various ROI size, especially the height, we were able to confirm that in Figure 7 the RMSE values did not decrease significantly even when we reduced the height to 64, which is more suitable for natural scene derived edges. Since the masking width of the adaptive window is automatically determined by edge angle and ROI height, no separate consideration for width was conducted in our paper. In cases with a certain level of noise (CNR = 25 and 5; Figure 7), the proposed method demonstrated significantly lower RMSE across the entire ROI height range (20-130) compared to the conventional method.

(a) CNR = inf (noise free)



Figure 4. Comparison of the MTFs measured by conventional and proposed methods at noise-free ROIs (a) and noisy ROIs (b). The first and second columns represent RMSEs between the measured and ideal MTF in ROIs with various edge sharpness (blur sigma from 0.5 to 1.2) and angles (from 1° to 45°). The last column representing conventional, proposed, and reference MTF graphs. RMSE, root mean square error; and MTF, modulation transfer function.



Figure 5. Comparison of LSFs and MTFs with the conventional fixed Tukey window and the proposed adaptive Tukey window.



Figure 7. Comparison of the RMSEs between the measured and ideal MTF (AVG±STD, n=10) according to ROI size. MTF, modulation transfer function; and RMSE, root mean square error.

Actual Test on Real Scene

Finally, we also evaluated performance in real natural-scene data. To examine the impact of noise, we swept the EIT from 4 milliseconds to 256 milliseconds. According to the result, including techniques for handling a broader angle range and enhancing robustness to noise, the proposed method detected a greater number of ROIs (marked as blue points; Figure 8) across all exposure times modified ROI size, as well as an expansion of the max angle range from 35° to 44°. Furthermore, we observed that the proposed method exhibits a reduced discrepancy between mean MTF and individual MTFs compared to the conventional method, and that mean MTF more effectively reflects overall trends of individual MTFs as shown in Figure 6. In fact, the MTF50/10 values also showed significantly smaller standard deviations in the proposed method compared to the conventional one as analyzed in Table 3. It

indicates that the proposed method represents higher consistency in sharpness with regardless of noise level.



Figure 8. Comparison of edges detected (upper row) by the existing (a) and proposed (b) NS-SFR methods and their corresponding MTFs (lower row) in images taken with various EITs (=16, 64, and 256ms). The noise level of CNR \geq 5 to emphasize the impact of the adaptive OSR/PS and window.

conventional and proposed methods with different EITS (n=4)						
	Conventional		Proposed			
	Mean	Standard Deviation	Mean	Standard Deviation		
MTF50	764.0	76.1	788.5	5.0		
MTF10	1233.8	28.1	1246.5	13.2		

Table3. Comparison of mean and standard deviation for the conventional and proposed methods with different EITs (n=4)

Conclusion

The NS-SFR enabled the evaluation of camera performance and image resolution and sharpness from general natural scenes, based on the existing e-SFR measurement method. However, the existing method had some limitations due to its limited edge angle range and vulnerability to noise. Therefore, our research aimed to overcome the limitations of the existing NS-SFR [3] by increasing the usable angle range and reducing the impact of noise in the process of NS-SFR. We demostrated that the proposed method enhanced its accuracy across a wider range of conditions while maintaining MTF consistency. Furthermore, by verifying the performance of our method in simulation and natural scene images, we verified that our method overcome the limitations of the conventional NS-SFR. These advancements further amplifies the potential of the NS-SFR, enabling multidimensional image quality measurement in the future.

References

- British Standard Institute, in *BS ISO 12233:2017 Photography -Electronic still picture imaging – Resolution and spatial frequency responses*, BSI Stand. Publ., 2017, p. 1–62.
- [2] British Standard Institute, in BS ISO 12233:2024 Digital cameras Resolution and spatial frequency responses, BSI Stand. Publ., 2024.
- [3] O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, "Estimation of ISO12233 Edge Spatial Frequency Response from Natural Scene Derived Step-Edge Data," *Journal of Imaging Science and Technology (JIST)*, vol. 65, no. 6, pp. 60402-1-60402-16(16), 2021.
- [4] O. van Zwanenberg, S. Triantaphillidou, A. Psarrou and R. B. Jenkin., "Analysis of natural scene derived spatial frequency responses for estimating camera ISO12233 slanted-edge performance (JIST-first)," in *Electronic Imaging*, 34, 2021.
- [5] D. Williams, "Benchmarking of the ISO 12233 slanted-edge spatial frequency response plug-in," in *IS&T PICS*, Springfield, VA, 1998.

- [6] K. Masaoka, "Edge-based modulation transfer function measurement method using a variable oversampling ratio," *Optics Express 29.23*, pp. 37628-37638, 2021.
- [7] K. Masaoka, "Practical edge-based modulation transfer function measurement," *Opt Express*, *27*, pp. 1345-1352, 2019.
- [8] Y. Wu, W. Xu, Y. Piao, W. Yue, "Analysis of edge method accuracy and practical multidirectional modulation transfer function measurement.," *Applied Sciences* 12.24: 12748, 2022.
- [9] O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, "Natural Scene Derived Camera Edge Spatial Frequency Response for Autonomous Vision Systems," in *IS&T/IoP London Imaging Meeting*, London, 2021.

Author Biography

Sara Lee received the B.S. and M.S. degrees in electronic engineering from Sogang University, Seoul, in 2019, and 2021, respectively. Her M.S. work focused on ultrasound imaging and signal processing. Since then she has worked in Samsung Electronics, Republic of Korea, as an engineer. Her work has focused on image sensor, image/video signal processing, computer vison, and image quality assessment.

Seungwan Jeon was born in Republic of Korea in 1989. He received the B.S. degree in biomedical engineering from Yonsei University in 2014, and the Ph.D. degree in creative IT engineering from POSTECH in 2020. His Ph.D. research focused on photoacoustic/ultrasound imaging techniques using image/signal processing, beamforming, and deep learning. Since 2020, he is with Samsung Electronics, Republic of Korea, as a Staff engineer. His current research interests include camera sensor, computer vision, and image quality assessment.

Subin Han received her B.S. degree in electrical engineering from Korea University in 2020. Since 2021, she has worked in Samsung Electronics, Republic of Korea, as an engineer. Her work has focused on computer vison, image/video signal processing, and image quality metric.

JOIN US AT THE NEXT EI!



Imaging across applications . . . Where industry and academia meet!





- SHORT COURSES EXHIBITS DEMONSTRATION SESSION PLENARY TALKS •
- INTERACTIVE PAPER SESSION SPECIAL EVENTS TECHNICAL SESSIONS •

www.electronicimaging.org

