A tool for deriving camera spatial frequency response from natural scenes (NS-SFR)

Van Zwanenberg, O., Triantaphillidou, S. and Jenkin, R.

A paper presented at Electronic Imaging 2023: Image Quality & System Performance XX, San Fransisco, 15 - 19 Jan 2023. The Society of Imaging Science and Technology.

Available from the publisher at:

https://doi.org/10.2352/EI.2023.35.8.IQSP-311 .

A Tool for Deriving Camera Spatial Frequency Response from Natural Scenes (NS-SFR)

Oliver van Zwanenberg, Sophie Triantaphillidou, Robin Jenkin; University of Westminster; London, United Kingdom

Abstract

Recent research on digital camera performance evaluation introduced the Natural Scene Spatial Frequency Response (NS-SFR) framework, shown to provide a comparable measure to the ISO12233 edge SFR (e-SFR) but derived outside laboratory conditions. The framework extracts step-edges captured from pictorial natural scenes to evaluate the camera SFR. Comprising two parts, the first utilizes the ISO12233 slanted-edge algorithm to produce an 'envelope' of NS-SFRs. The second estimates the system e-SFR from this NS-SFR data.

One current drawback of this proposed methodology has been the computation time. Although successful in e-SFR estimation, the process was not optimized; it derived NS-SFRs from all suitable step-edges before statistically treating the results to estimate the e-SFR.

This paper presents changes to the framework processes, aiming to optimize the computation time for real-world implementation while maintaining e-SFR estimation accuracy. The developments include an improved framework structure and an edge isolation step that is easier to compute. The resulting code has been incorporated into a self-executable user-interface prototype, available on GitHub, allowing users to select images for e-SFR camera estimation, for different radial annuli and camera orientations.

Introduction

Previous publications [1–3] presented research that developed a novel automated methodology that estimates the standard edge Spatial Frequency Response (e-SFR) directly from pictorial natural scenes. The proposed methodology [1, 2] was split into two parts. The first part located, isolated and then verified step-edges from photographed scenes. These step-edges were put through the ISO12233:2017 e-SFR algorithm, providing Natural Scene derived Spatial Frequency Responses (NS-SFRs). The second part applied thresholds to the NS-SFRs to find the edges that were most likely to return test chart response for estimating the standard e-SFR. The resulting e-SFR estimates were derived for six radial annuli across the field of view and the entire frame [1–3]. This workflow is depicted in Figure 1.

Investigations were conducted using three large, diverse image datasets, comprising 1-2K images each, captured for the purpose [3], using two near-linear digital single-lens reflex (DSLR) cameras and one highly non-linear smartphone system. The resulting e-SFRs were shown to be within the limits of the ISO12233 e-SFRs derived in the lab, from test chart inputs, for all the (near-)linear systems, i.e., the estimates stayed within +/- one standard deviation of the equivalent test chart measurement. Results from the highly non-linear system showed signs of scene-and-processing dependency (SPD); this SPD behavior has not yet been quantified. Overall, the results suggest that the proposed method is a viable alternative to the ISO technique for characterizing linear and near-linear camera systems, and hints at scene-dependent image processes in the non-linear system.

The measuring framework can be used in several applications. A live-SFR measurement is of significant interest, potentially allowing a camera system to be monitored to ensure it is fully operational for decision-critical tasks. Live-SFRs can open new possibilities in the evaluation of autonomous vision systems, such as autonomous vehicles [4], CCTV and security systems. Another useful application is the ability to obtain SPD-SFRs for characterizing highly non-linear adaptive Image Signal Processing (ISP), which is now an integral part of modern camera systems. Such systems have become non-linear black boxes, where linear and non-linear processes are inseparable and thus difficult to fully characterize. NS-SFR analysis has the potential to provide further insight into non-linear and scene-dependent behavior [1].

One drawback of the NS-SFR framework has been the computation time since priority was given to measurement accuracy rather than efficiency. The average computation time per image (calculated from 1500 RAW (green channel) 36.3-megapixel (MP) images) was 16 minutes. The framework first derives NS-SFRs from all suitable step-edges and further validates and statistically treats the results to estimate the e-SFR. This two-part process was shown to be highly inefficient since it resulted in only 3.4% of the isolated NS-SFRs being utilized in the e-SFR estimate.

This paper presents changes to the framework processes, aiming to optimize the computation time for real-world implementation while maintaining e-SFR estimation accuracy. The rationale behind the significant optimizations made to the framework is described. The computational improvements and the error of the e-SFR estimates are presented and compared to the previous workflow for a standard test-chart and the datasets presented in the original publications [1, 2]. Finally, the updated measurement framework was packaged in an open-source, self-executable application and is readily available to download via GitHub for SFR camera evaluation.

Workflow Optimizations

The primary step in optimizing the estimation of the e-SFR from natural pictorial images was to reorder the processes so that the step-edges were fully validated before they underwent heavy computational processing. This resulted in 96.6% of unutilized edges being deselected as early as possible. The flowchart of the previous iteration and the optimized workflow are compared in Figure 1. Further, the code was optimized to perform the same tasks with lower computational requirements.

In detail, first, rather than re-evaluating each edge multiple times, the step-edge verification, contrast deselection and edge-of-interest masking processes were combined into one process. The majority of edges, which were unsuitable for e-SFR estimation, were deselected, eliminating repetitive calculations. The edge contrast selection at an early stage in the workflow resulted in 66% less processing.

Secondly, an alternative to pixel stretching, a filter previously developed as part of the framework to isolate the edge-of-interest that removed unwanted scene textures and background objects, was established. In this, the edge isolation is achieved using an edge mask, extracting the region where the actual step-edge resides, see Figure 2.

In the revised framework, once the edges are cropped from the image in regions of interest (ROIs), the NS-SFRs are measured using *sfrmat4* [5] (*ISO12233:2017* [6]). The NS-SFR measurements follow the methodology stated in the previous publications [1, 2]. Only the luminance e-SFRs are measured; RGB e-SFRs require 4x the processing time (for all 3 color channels and the luminance).

To obtain the 'sharpest' edges, i.e., the edges most likely to return a test chart response, the previous iteration of the workflow took the 10^{th} percentile of the distribution of the Line Spread Function (LSF) Full Width Half Maximum (FWHM) per radial annuli. This was an effective method to estimate the system e-SFR. However, it also deselects 90% of the already processed NS-SFRs. The optimized workflow has a stricter edge verification process, resulting in fewer but 'sharper' step-edges being processed. As a result, a 'sharpness' threshold is no longer required; all selected step-edges contribute to the e-SFR estimation.

Previously, a multidimensional coordinate system was used to interpolate the NS-SFRs across multiple parameters (i.e. angle, contrast, noise). This is a valuable tool for NS-SFR analysis, removing anomalous data across parameters. However, for the e-SFR estimation, it can result in additional unrequired processing. Therefore, in this optimized workflow, the multidimensional

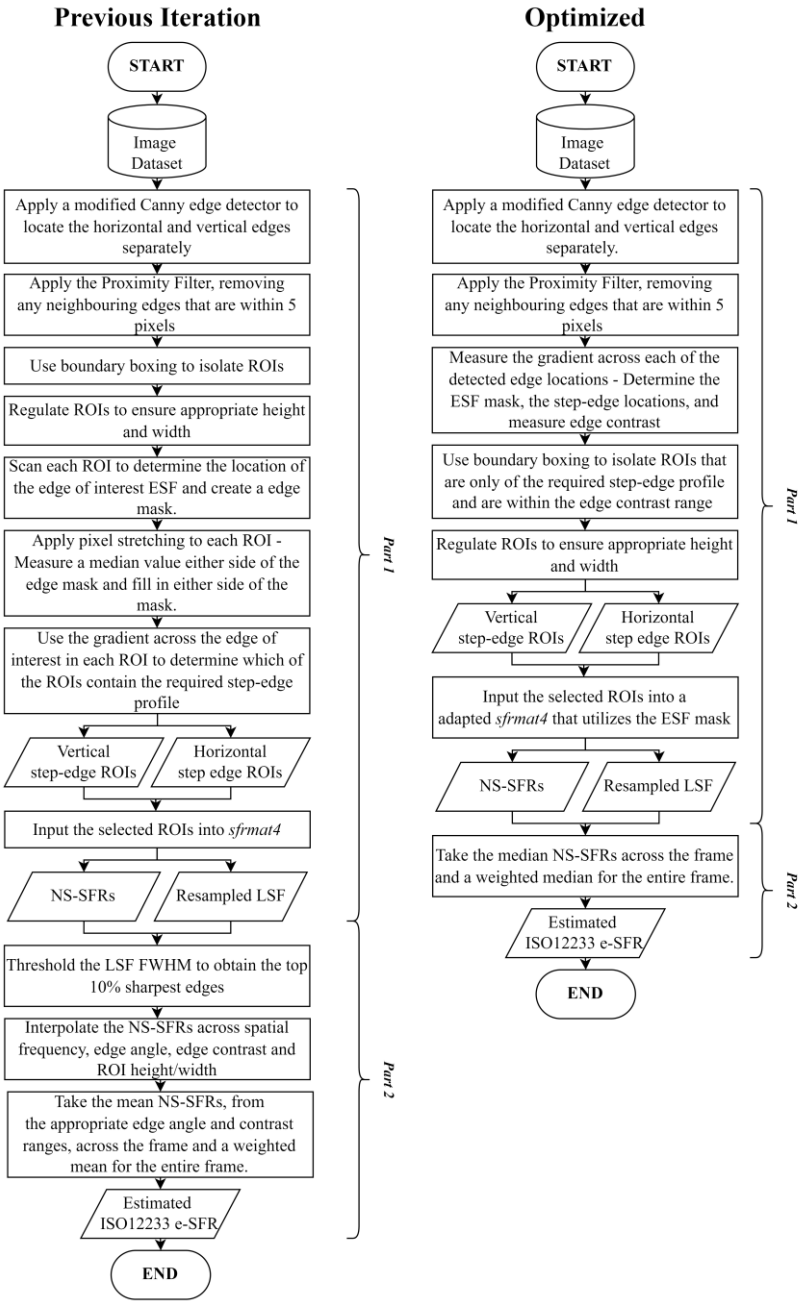

Figure 1. *Comparison between the structure of the previous iteration and optimized workflows.*

coordinate system was replaced by calculating the median NS-SFRs per radial annuli and a global weighted median NS-SFR across the field of view.

Step-edge Isolation

Influenced by the filtered tails procedure [7], pixel stretching was established for step-edge isolation and worked consistently to remove unwanted natural scene textures, reducing the bias of image noise and diminishing the error caused by non-uniformity [1, 2].

Although effective, pixel stretching is computationally heavy and, with thousands of ROIs, adds significant runtime to the e-SFR estimation. An alternative solution was developed that uses edge detection (horizontal and vertical gradient direction) to establish the abovementioned edge mask, efficiently combining the two processes. The SFR algorithm (*sfrmat4* [5]) was adjusted to utilize this edge mask directly, windowing the LSF in each ROI row before resampling the LSF down the slope of the edge-of-interest.

The derived e-SFR from both edge isolation methods are identical (Figure 2a) and present the same advantages of noise and non-uniformity reduction. The step-edge isolation also improves results in scenarios where pixel stretching caused anomalous SFRs, for example, when color fringing is present in the pictorial natural scene ROI (Figure 2b and 2c).

Results

e-SFR Estimation

The performance of the optimized workflow was evaluated using two characterized datasets of 1500 images from two different camera systems:

i. **DSLR** – the green channel of the RAW images of the Nikon D800 DSLR camera with the Nikkor 24-70mm *f*/2.8G ED Lens, set at 24mm *f*/4. The sensor is 36.3MP with a pixel pitch of 4.87µm.

ii. **Smartphone** – images captured using the Apple iPhone 7 rear camera (4mm *f*/1.8). The sensor is 12MP with a pixel pitch of 1.22µm.

For both systems, images were used to estimate the e-SFR across the frame. The DSLR system is considered (near-)linear; the evaluation used the green channel sensor image data. The smartphone is highly non-linear, where the luminance data of the compressed demosaiced image output was used in the evaluation. The absolute error between these estimates and the standard test chart measurements (*ISO12233:2017* [6]) was calculated and compared to the previous iteration of the workflow.

To obtain system e-SFR estimates across the frame, the field of view was divided into radial annuli. Previous works [1–3] used an even radial distance to segment the camera frame. However, this method resulted in an uneven frame area per annulus, meaning the part-way regions had the potential to contain more edges than the center and corners of the frame [1]. The data in this publication was calculated with adjusted segmentation to distribute the radial annuli frame area equally, providing a higher probability for each annulus to contain a similar number of edges for the e-SFR estimation.

The number of annuli has also been changed. In the previous publications, six segments were used to estimate annular e-SFRs and provide a weighted global average of the entire frame. Whereas now, the number has been reduced to three. This reduction of the number of annuli improved the estimation accuracy by deriving e-SFRs from a larger amount of edge data. This was a necessary step because, through optimizing the workflow, the step-edge selection resulted in fewer edges than in the previous iteration.

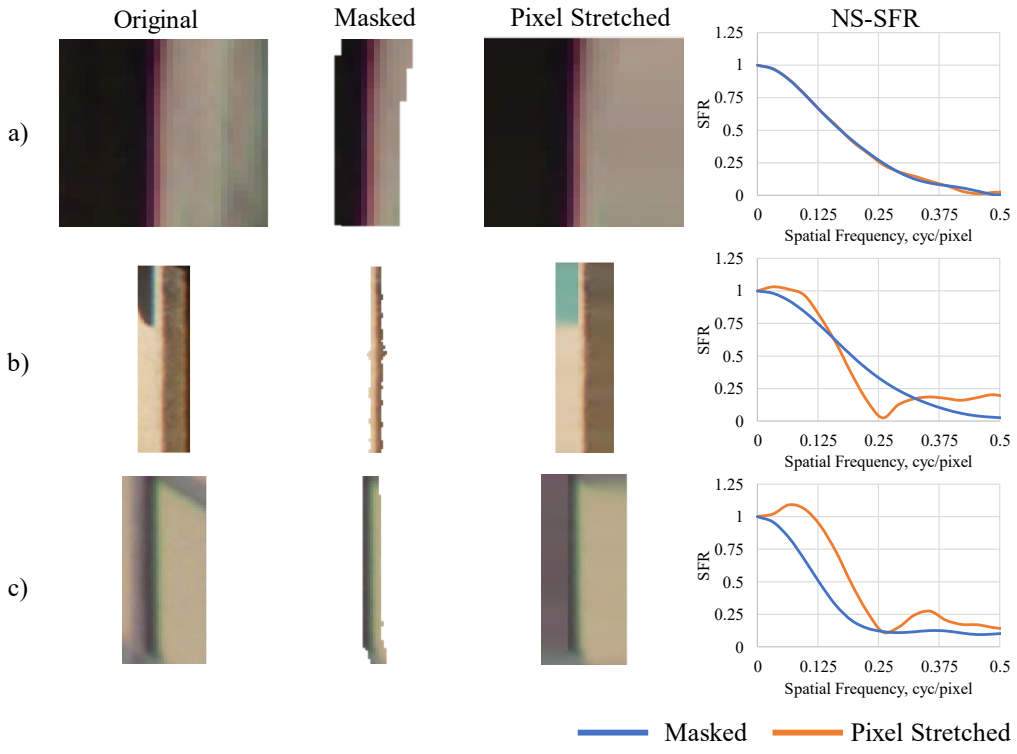

***Figure 2**. Comparing the NS-SFRs that result from pixel stretching and optimized edge-masked isolation on three different pictorial natural scene ROIs. Under certain circumstances, the Masked ROI provided better results compared to the Pixel-Stretched ROI.*

Figure 3 contains the vertical e-SFR estimates for the DSLR system. The optimized workflow is directly compared to the previous iteration of the framework in relation to the ISO12233 e-SFR test chart measurements. Horizontal e-SFRs show similar trends. The figure also contains the absolute error between the estimated system e-SFR, derived from the median of the ISO12233 e-SFR. Similarly, with results from the previous iteration, it is observed that the optimized framework provides estimated system e-SFRs that lay within, or close to (Radial Annuli 1/3), +/- one standard deviation of the test chart measurements. A high-frequency bias remains in the e-SFR estimations, caused by the ROIs containing image noise and high-frequency scene textures [2, 4].

Table 1 contains the number of edges isolated from the DSLR dataset compared to the number of edges utilized in the e-SFR estimation. The previous iteration of the workflow calculated the e-SFR from a significantly greater number of edges than the optimized framework, which would account for the slight discrepancy in estimation between the two iterations of the workflow.

Figure 4 contains the weighted global average system e-SFR estimation for the smartphone system, comparing the optimized and previous iteration of the framework in relation to the ISO12233 e-SFR test chart measurement. NS-SFR measurements from highly non-linear systems are SPD [1, 3]. The optimized workflow also shows scene dependency originating from non-linear processing. When isolating step-edges from test charts, adaptive processing has an insignificant effect; chart edges are preserved and enhanced, resulting in a low-frequency boost in the e-SFR. In contrast, sharpening step-edges in complex natural scenes does not result in a low-frequency e-SFR boost, as the inclusion of surrounding scene content and textures reduces sharpening effects on natural scene edges [3]. In this case, the e-SFR derived from natural scenes is a more representative camera performance measure than the e-SFR derived from test charts.

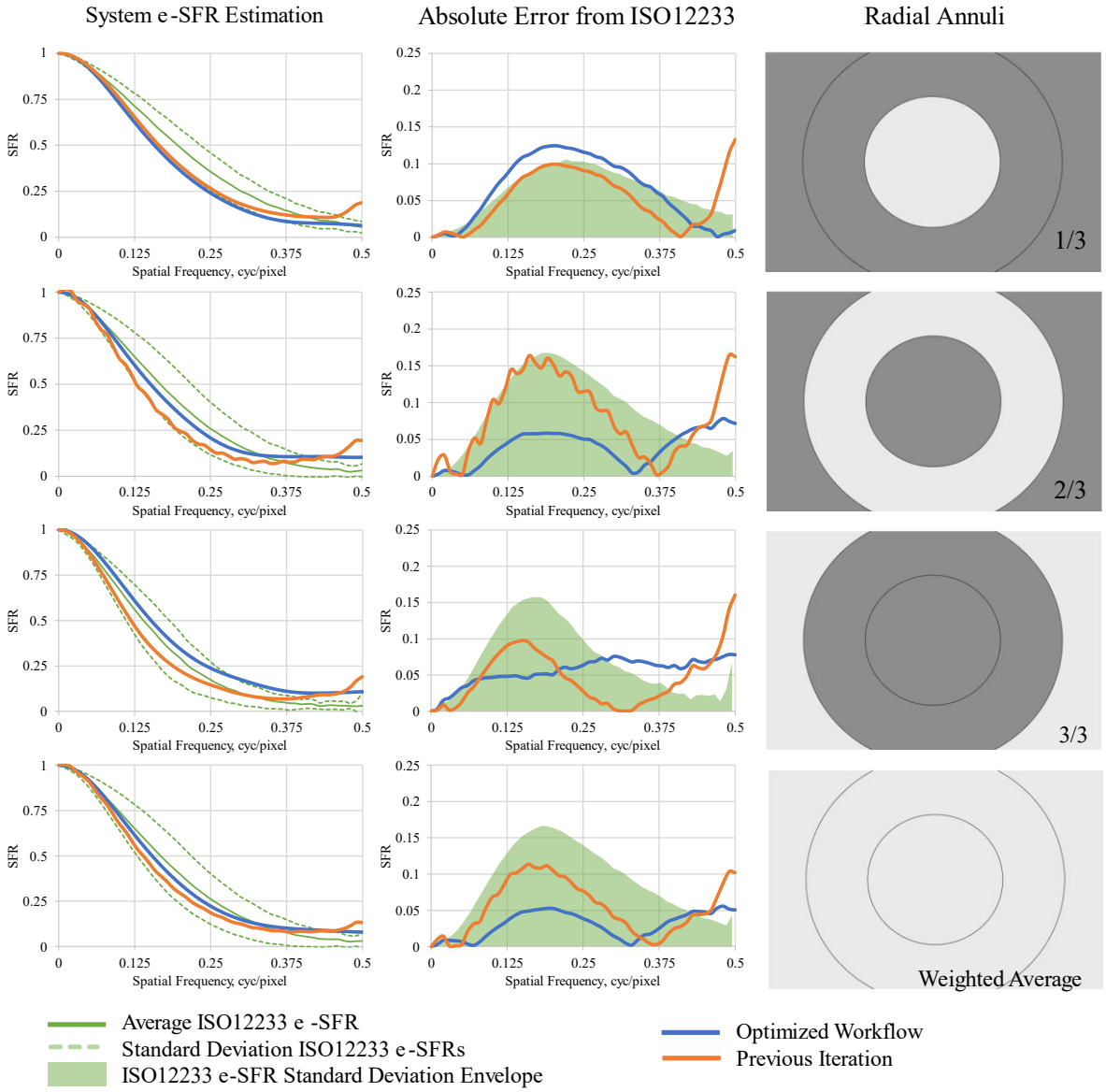

***Figure 3**. DSLR vertical system e-SFR estimation for three radial annuli and a weighted average of all annuli. The first column contains the estimated system e-SFR in relation to the ISO12233 e-SFR. The second column contains the absolute error between the estimated system e-SFR from the median of the ISO12233 e-SFR. The third column contains a visual representation of the radial annuli from which the data belongs.*

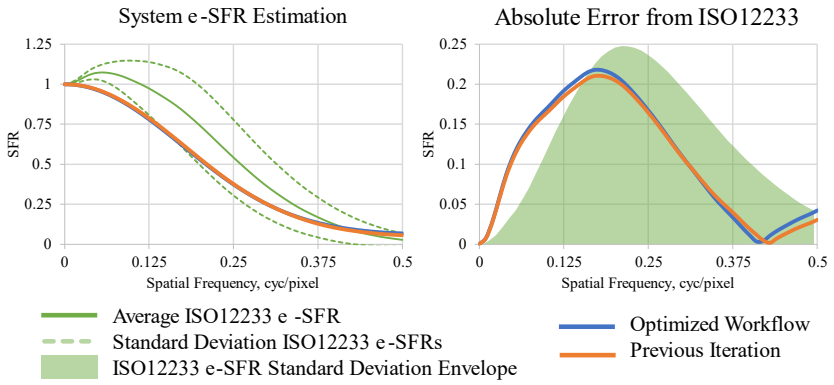

***Figure 4**. Smartphone vertical system e-SFR estimation for the weighted average of all three annuli. The first column contains the estimated system e-SFR in relation to the ISO12233 e-SFR. The second column contains the absolute error between the estimated system e-SFR from the median of the ISO12233 e-SFR.*

		Optimized	Previous Iteration
Edges Isolated	Annuli 1	67	39852
	Annuli 2	41	28191
	Annuli 3	24	18915
Edges Utilized	Annuli 1	67	4100
	Annuli 2	41	2936
	Annuli 3	24	1898
Utilization (%)	Annuli 1	100	10.3
	Annuli 2	100	10.4
	Annuli 3	100	10

***Table 1.** This table contains the number of horizontal edges isolated from the DSLR dataset using the optimized and previous iteration of the workflow, and the number of these edges used in the vertical system e-SFR estimation, shown as the number of edges and as a percentage of the isolated ROIs.*

Computation Improvement

The framework is in two parts. The first measures the NS-SFRs, and the second estimates the system e-SFR. There have been significant computational savings for both parts in the new optimized workflow.

Part 1

The computational time for the first part of the framework improved by a factor of 22. For instance, the optimized computation time to process the DSLR images took an average of 0.73 minutes per image, while the previous iteration for the same images took an average of 16 minutes per image. The computational time to process the entire dataset (all 1500 images) using a 16-core Central Processing Unit (CPU) that ran at 4.2GHz, was 68 minutes, compared to the previous 1470 minutes (24.5 hours).

Part 2

The second part of the framework was improved by a factor of 2. The e-SFR estimation took 10 minutes for the DSLR dataset compared to 20 minutes in the previous workflow iteration.

Live-SFR

The workflow cannot yet be implemented with live systems. Further optimization is required. For the DSLR system, the computational time will need to be improved by a factor of 1314 to achieve a live-SFR measurement (30 frames per second). That said, many live systems would have a much lower pixel resolution than the 36.3MP DSLR sensor, significantly impacting the computation time per frame/image. The estimation from the smartphone 12MP system is approximately twice as fast as the higher-resolution DSLR.

Self-Executable User Interface

This optimized code has been packaged in a self-executable Graphical User Interface (GUI), shown in Figure 5. It is available on GitHub: *https://github.com/OliervZ11/NSSFR-GUI*

***Figure 5**. Screen Capture of the prototyped Self-Executable GUI.*

This GUI provides a means to derive camera e-SFRs by processing camera datasets, with control over the contrast, angle, and ROI thresholds, as well as upload already processed data and display the results. ISO12233 test chart data can be processed to compare results with e-SFR estimations. Results are exported as .xlsx files. The application also provides a modelled MTF based on the sensor pixel pitch and lens *f*/number [8].

The unoptimized workflow is also available through this GUI. It has been provided for SPD-SFR evaluation, outputting the NS-SFRs, ROI image coordinates, edge contrast and edge angle in a .xlsx file per image (see discussions). For the purpose, the unoptimized workflow has been adapted to use masked edge isolation instead of pixel stretching.

Two additional tools that have not been included in the GUI are included as MATLAB files in the GitHub repository:

1) a GUI that allows the selection of edges from a test chart and structures the e-SFR results in a .xlsx file to compare directly with the estimated system e-SFRs per radial annuli.

2) A tool to label and apply scene classification to a dataset of images. Two scene classification tools are provided. The first is an automated method which is based on a modified version of AlexNet [9] to classify the general environment of natural scenes (Man-made, Indoor and Nature [1, 3]). The second is a manual method of scene classification, using a GUI to cycle through a dataset of

images, allowing the user to label the main light source, the primary image subject and the general environment (Man-made, Indoor or Nature).

Discussion

The NS-SFR framework workflow and implementation have been successfully optimized, improving the computational time required to derive e-SFRs from natural scene images by a factor of 22, while keeping the accuracy consistent. Hardware optimization can further speed up the workflow to achieve a live-SFR measurement by employing Graphics Processing Unit (GPU) acceleration.

The original workflow version gathered more data, which was then culled down, and averaged the remainder to provide an estimated system e-SFR. In comparison, the optimized workflow gathered fewer step-edges but with suitable edge contrast, angle and ROI window size.

With a 'stricter' step-edge verification process, the optimized workflow removed the requirement for an edge 'sharpness' threshold. Previously, the narrowest 10% of LSFs were selected for the system e-SFR estimation, as they would most likely return a test chart response. The removal of this threshold has meant that all processed NS-SFRs are used in the e-SFR estimation. However, it also introduces uncertainty, as an isolated edge may have the correct edge profile but a low gradient, which may lower the e-SFR estimation. Future iterations of the workflow should measure 'sharpness', perhaps measuring gradients at the edge detection/step-edge verification stage.

Despite the universal use of the SFR by industry and standard bodies for characterizing camera system performance, to date, no method successfully derives SFRs from non-linear camera systems, which represent the large majority of digital cameras. Further research should study camera system performance variations originating from different SPD non-linear processes and link it statistically with original scene content variations. Examining contents surrounding the edges used to derive NS-SFRs and linking them to different camera/ISP components may produce more representative camera performance measure(s) than the current SFR, depicting real-world performance while being compatible with current industry standards. As the workflow was optimized with the estimation of the ISO12233 e-SFR in mind, the unoptimized workflow has also been incorporated in the e-SFR estimation GUI. It returns a much larger number of NS-SFRs that can be used for SPD SFR studies. In addition, scene classification tools have been provided in the GitHub repository for use in such studies.

The next iteration of the ISO12233 standard is anticipated in 2023. This latest standard will come with many improvements to the ISO12233 to increase the accuracy and precision of the measured e-SFR. Once released, these advances in the ISO12233 [10] would benefit future revisions of the NS-SFR and system e-SFR estimation methodology. Such advantages would include the following:

- 5th order polynomial fitting function
- Edge angle variation correction
- Non-uniform illumination correction

Conclusion

In summary, a previously established framework workflow used to estimate the system e-SFR from natural scene images was optimized to reduce processing time while keeping accuracy at the same levels. Using a dataset of images from a near-linear system, we found that the estimated system e-SFR stays within +/- one standard deviation of an equivalent test chart e-SFR measurement, whilst the processing time is 22x faster than the original version. Non-linear system e-SFR estimation from a smartphone camera image dataset requires further investigation, but showed that the method provides opportunities for a scene-dependent camera evaluation.

The framework processes are contained in a self-executable prototype. The application returns median global system e-SFR estimation from across the entire frame, as well as localized estimated e-SFRs for selected radial annuli and camera orientations. This is done by simply uploading camera datasets and selecting the desired measurement functions. GPU acceleration can optimize the processes further to return e-SFR measurements from live imagery. Finally, the application also provides means to extract NS-SFR envelopes from natural scenes that can facilitate the investigation of SPD camera performance.

References

[1] O. van Zwanenberg, "Camera Spatial Frequency Response Derived from Pictorial Natural Scenes," University of Westminster, 2022.

[2] O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Estimation of ISO12233 Edge Spatial Frequency Response from Natural Scene Derived Step-Edge Data," *Journal of Imaging Science and Technology*, vol. 65, no. 5, pp. 60402-1-60402–16, 2022.

[3] 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," *Journal of Imaging Science and Technology*, vol. 65, no. 6, pp. 60405-1-60405–12, 2022.

[4] O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Natural Scene Derived Camera Edge Spatial Frequency Response for Autonomous Vision Systems," in *London Imaging Meeting*, 2021, pp. 88–92, doi: https://doi.org/10.34737/v644w.

[5] P. D. Burns, "sfrmat," *Burns Digital Imaging*, 2019. http://burnsdigitalimaging.com/software/sfrmat/ (accessed Dec. 01, 2022).

[6] British Standard Institute, "BS ISO 12233:2017 Photography - Electronic still picture imaging - Resolution and spatial frequency responses," *BSI Standards Publication*, pp. 1–62, 2017.

[7] D. Williams and P. D. Burns, "Evolution of Slanted Edge Gradient SFR Measurement," *Proc. SPIE 9016 - Image Quality and System Performance XI*, vol. 9016, p. 8, 2014, doi: 10.1117/12.2042706.

[8] R. Jenkin and P. Kane, "Fundamental Imaging System Analysis for Autonomous Vehicles," in *Autonomous Vehicles and Machines Conference*, 2018, vol. 2018, no. 17, pp. 105-1-105–10, doi: 10.2352/issn.2470-1173.2018.17.avm-105.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in neural information processing systems, vol. 1, pp. 1097–1105, 2012.

[10] P. D. Burns, K. Masaoka, K. Parulski, and D. Wueller, "Updated Camera Spatial Frequency Response for ISO 12233," in *Electronic Imaging 2022, Image Quality and System Performance XIX*, 2022, pp. 1–7.

Author Biography

Oliver van Zwanenberg received his BSc at the University of Westminster, London, in 2017. Then moved on to pursue his PhD in that same year. His PhD, titled 'Camera Spatial Frequency Response Derived from Pictorial Natural Scenes', was awarded in 2022. In his thesis, a methodology was established that adapted the standard ISO12233 e-SFR to utilize captures of natural scenes to estimate the camera system performance. He is currently working as a research associate at the University of Westminster, and is a keen photographer and videographer.