

WestminsterResearch

<http://www.westminster.ac.uk/westminsterresearch>

**Camera Spatial Frequency Response Derived from Pictorial
Natural Scenes**

Van Zwanenberg, O.

A PhD thesis awarded by the University of Westminster.

© Dr Oliver Van Zwanenberg, 2022.

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Camera Spatial Frequency Response Derived from Pictorial Natural Scenes

Oliver van Zwanenberg, B.Sc. (Hons)

School of Computer Science and Engineering

College of Design, Creative and Digital Industries

University of Westminster

A thesis submitted in partial fulfilment of the requirements of the University of
Westminster for the Degree of Doctor of Philosophy.

This research programme was completed within the Computational Vision and Imaging
Technology Research Group.

Abstract

Camera system performance is a prominent part of many aspects of imaging science and computer vision. There are many aspects to camera performance that determines how accurately the image represents the scene, including measurements of colour accuracy, tone reproduction, geometric distortions, and image noise evaluation. The research conducted in this thesis focuses on the Modulation Transfer Function (MTF), a widely used camera performance measurement employed to describe resolution and sharpness. Traditionally measured under controlled conditions with characterised test charts, the MTF is a measurement restricted to laboratory settings. The MTF is based on linear system theory, meaning the input to output must follow a straightforward correlation. Established methods for measuring the camera system MTF include the ISO12233:2017 for measuring the edge-based Spatial Frequency Response (e-SFR), a sister measure of the MTF designed for measuring discrete systems.

Many modern camera systems incorporate non-linear, highly adaptive image signal processing (ISP) to improve image quality. As a result, system performance becomes scene and processing dependant, adapting to the scene contents captured by the camera. Established test chart based MTF/SFR methods do not describe this adaptive nature; they only provide the response of the camera to a test chart signal. Further, with the increased use of Deep Neural Networks (DNN) for image recognition tasks and autonomous vision systems, there is an increased need for monitoring system performance outside laboratory conditions in real-time, i.e. live-MTF. Such measurements would assist in monitoring the camera systems to ensure they are fully operational for decision critical tasks.

This thesis presents research conducted to develop a novel automated methodology that estimates the standard e-SFR directly from pictorial natural scenes. This methodology has the potential to produce scene dependant and real-time camera system performance measurements, opening new possibilities in imaging science and allowing live monitoring/calibration of systems for autonomous computer vision applications.

The proposed methodology incorporates many well-established image processes, as well as others developed for specific purposes. It is presented in two parts.

Firstly, the Natural Scene derived SFR (NS-SFR) are obtained from isolated captured scene step-edges, after verifying that these edges have the correct profile for implementing into the slanted-edge algorithm. The resulting NS-SFRs are shown to be a function of both camera system performance and scene contents. The second part of the methodology uses a series of derived NS-SFRs to estimate the system e-SFR, as per the ISO12233 standard. This is achieved by applying a sequence of thresholds to segment the most likely data corresponding to the system performance. These thresholds a) group the expected optical performance variation across the imaging circle within radial distance segments, b) obtain the highest performance NS-SFRs per segment and c) select the NS-SFRs with input edge and region of interest (ROI) parameter ranges shown to introduce minimal e-SFR variation. The selected NS-SFRs are averaged per radial segment to estimate system e-SFRs across the field of view. A weighted average of these estimates provides an overall system performance estimation.

This methodology is implemented for e-SFR estimation of three characterised camera systems, two near-linear and one highly non-linear. Investigations are conducted using large, diverse image datasets as well as restricting scene content and the number of images used for the estimation. The resulting estimates are comparable to ISO12233 e-SFRs derived from test chart inputs for the near-linear systems. Overall estimate stays within one standard deviation of the equivalent test chart measurement. Results from the highly non-linear system indicate scene and processing dependency, potentially leading to a more representative SFR measure than the current chart-based approaches for such systems. These results suggest that the proposed method is a viable alternative to the ISO technique.

Contents

Abstract	i
Contents	iii
List of Figures	vii
List of Tables.....	xiii
Acknowledgements	xiv
Author's Declaration.....	xv
Chapter 1 Introduction	1
1.1 Camera System Performance	1
1.2 The Limitations of Current MTF Approaches	4
1.3 Scope of Work.....	4
1.4 Thesis Structure.....	5
1.5 Contributions to Knowledge	7
1.6 Related Work	7
1.6.1 Publications	7
1.6.2 Presentations at Conferences and Symposia	8
1.6.3 Awards	10
Chapter 2 Background to System Performance	11
2.1 The Modulation Transfer Function	11
2.1.1 System Requirements	11
2.1.2 The Sinewave Method.....	12
2.1.3 The Edge Method	13
2.1.4 The Texture Method.....	15
2.1.5 Results and Implementation of the MTF Measurements	16
2.2 Digital Image Formation	16
2.2.1 Imaging Equation	16
2.2.2 Digital Sampling	17
2.2.3 Expected System MTF	19
2.3 The Slanted-Edge Method.....	22
2.3.1 System Linearisation	22
2.3.2 The Slanted-Edge Algorithm	24
2.3.3 Benefits of the Slanted-Edge Method	26
2.4 Sources of e-SFR Variation	27

2.4.1	Edge and ROI Parameter Ranges	27
2.4.2	Non-Linear Image Signal Processing	29
2.4.3	Optical Performance Variation	30
2.5	Sources of e-SFR Error	31
2.5.1	Image Noise Corruption	31
2.5.2	Tonal Non-Uniformity	34
2.5.3	Signal Clipping (Saturation)	35
2.6	Assessing e-SFRs	36
2.6.1	Comparison Metrics	36
2.6.2	SFR Orientation	38
2.7	Improvements to ISO 12233	40
2.7.1	General e-SFR Developments	40
2.7.2	Real-Time MTF	42
2.8	Natural Scene Derived System Performance	44
2.8.1	Edge Isolation	44
2.8.2	Signal-and-Process-Dependent MTFs	45
2.8.3	Blind Deconvolution	45
2.8.4	Neural Network MTF Estimation	46
2.9	Natural Scene Derived SFR Considerations	48
2.10	Image Noise	51
2.11	Summary	53
Chapter 3	Spatial Frequency Response Derived from Natural Scene Images	55
3.1	Overview of the NS-SFR Framework	55
3.2	Edge Detection	58
3.2.1	Canny Edge Detection (Workflow 1)	58
3.2.2	Matched Filter (Workflow 2)	59
3.2.3	Comparison between Edge Detection Workflows	61
3.3	Region of Interest Selection, Isolation and Verification	62
3.3.1	Region of Interest Selection	62
3.3.2	Edge Isolation – Pixel Stretching	66
3.3.3	Step-Edge Validation	73
3.4	Removal of Anomalous ROIs	76
3.4.1	Non-Uniformity Streaking Detection	76
3.4.2	Multidirectional Edges Detection	77
3.5	The NS-SFR Measure	78
3.6	Initial Testing: The Image NS-SFR Envelope	79
3.7	Summary	83

Chapter 4	Sources of Variation in the Edge-Based Spatial Frequency Response	85
4.1	Edge and ROI Parameter Variation.....	85
4.2	Variation Study Methodology.....	86
4.2.1	Simulated Step-Edges	87
4.2.2	Step-Edge Test Chart Captures	89
4.2.3	Applying Image Noise	90
4.2.4	Variation Assessment.....	93
4.3	Multi-Edge Parameter Variation Results	94
4.3.1	Simulated e-SFR Results	94
4.3.2	Captured Test Chart Results.....	97
4.4	Parameter Selection for System e-SFR Estimation.....	100
4.5	Summary	101
Chapter 5	Estimation of the System Spatial Frequency Response	103
5.1	System e-SFR Estimation Methodology.....	103
5.1.1	Overview of the Methodology	103
5.1.2	Data Organisation.....	104
5.1.3	Selection of Suitable Step-Edges	106
5.1.4	Data Binning and System e-SFR Estimation	108
5.2	Validation of the Established Parameter Ranges	109
5.3	System e-SFR Estimation Results.....	113
5.3.1	Characterised Systems and the Respective Image Datasets.....	113
5.3.2	Region of Interest Utilisation	116
5.3.3	DSLR (Near-Linear) Camera System e-SFR Estimations	119
5.3.4	Smartphone (Highly Non-Linear) Camera System.....	131
5.4	Summary	135
Chapter 6	Evaluation of Image Dataset Size and Scene Classification for System e-SFR Estimation	137
6.1	Dataset Size Evaluation Methodology.....	137
6.1.1	Camera System under Evaluation	137
6.1.2	Scene Classification	139
6.1.3	Subdividing the Dataset	143
6.2	System e-SFR Estimation Results from Subdivided Dataset.....	144
6.3	Summary	149
Chapter 7	Discussion	151
7.1	The Proposed Methodology for e-SFR Measurement	151
7.1.1	Estimated System e-SFR Methodology	152

7.1.2	Methodology Improvements	155
7.2	Implementation	157
7.2.1	Code Optimisation to Improve Usability	157
7.2.2	Recommended Settings and Thresholds	158
7.2.3	Improvements to the ISO 12233 Algorithm	159
7.2.4	Scene Dependant e-SFRs	161
7.3	Comparison to the PSF CNN Approach	162
7.4	Potential Applications	166
7.4.1	Camera Performance Comparison	166
7.4.2	Image Quality Metrics.....	166
7.4.3	Autonomous Vision Systems	167
Chapter 8	Conclusions and Recommendations	169
8.1	Conclusions	169
8.2	Further Work.....	171
Appendix A	The ISO12233 e-SFR Measurement	174
Appendix B	E-SFR Estimation Summary Metric.....	176
Appendix C	All Vertical e-SFR Estimation Results	180
Appendix D	Tangential DSLR e-SFR Estimation	190
Appendix E	MATLAB Code	192
Abbreviations	198
References	200

List of Figures

Figure 1.1 Examples of a low-performance MTF, a high-performance MTF and a MTF from a sharpened edge.	1
Figure 1.2 Visualisation of a simplified digital camera system pipeline.	2
Figure 2.1 Test chart input patterns for MTF and SFR measurements.	11
Figure 2.2 Illustration of a perfect step-edge, containing maximum modulation across all spatial frequencies.	13
Figure 2.3 MTF plots depicting the frequency content of a perfect step-edge and a near-perfect step-edge, with respect to an example of a system MTF. The shaded area above the example system MTF illustrates the region where the frequency content of an input step-edge is <i>suitable</i> for measuring the MTF but requires correction. The shaded area below the system MTF illustrates the frequency content <i>unsuitable</i> for measuring the MTF.	14
Figure 2.4 The fundamental relationships between spread functions and system performance, adapted from [7, p. 133].	17
Figure 2.5 Modelled in-phase (max), out-of-phase (min) and average MTF_{sys} for the <i>Nikon D800</i> DSLR system up to four times the Nyquist limit.	20
Figure 2.6 The maximum and minimum a) LSF and b) ESF for the simulated <i>Nikon D800</i> DSLR system. The spread of the system is marked with vertical broken lines, depicting the point at which the LSFs reaches 50% of the maximum peak (FWHM).	21
Figure 2.7 The ISO12233 slanted-edge method flowchart, adapted from the ISO12233:2017 [5].	23
Figure 2.8 Registering scan lines with a slanted edge, allowing the ESF to be resampled. Adapted from [22].	26
Figure 2.9 The mean oversampling factor, capped at 8-times, for simulated step-edges ranging from 0 to 45 degrees. The ROI window height used was 30 pixels. Adopted from [67].	28
Figure 2.10 Illustration of the e-SFR variation introduced by edge location within the field of view, captured using the <i>Nikon D800</i> with a 24mm lens at $f/4$ [76].	31
Figure 2.11 The output e-SFRs for eight simulated ROIs with decreasing SNR values.	32
Figure 2.12 The confidence limits depicting the possible spread due to noise corruption from a) the OTF and b) the MTF. Adopted from [39].	33
Figure 2.13 The bias introduced to the e-SFR with low frequency non-uniformity. The ESF and SFR demonstrate how these non-uniformities affect the result compared to a uniform ROI. Adapted from Imatest [77].	35
Figure 2.14 Simulated noiseless ROIs at a 0.60 Michelson Contrast with mid-range unclipped pixel values, soft-clipped pixel values, and hard-clipped pixel values. The output ESFs and the resulting e-SFR are plotted, illustrating the clipping error.	36
Figure 2.15 Illustration of the edge orientations for sagittal and tangential e-SFR analysis in relation to the optical imaging circle and field of view.	38

Figure 2.16 The process of plotting the sagittal and tangential e-SFRs/MTFs for 25%/50%/75% of the Nyquist frequency and how it relates to the traditional MTF bell curve. Adapted from [15].	39
Figure 2.17 A visual representation of the reverse projection method compared to the standard forward projection method. Adopted from [64].	41
Figure 2.18 Selection process for Masaoka's real-time MTF measurement. Adapted from [66].	43
Figure 2.19 Four examples of implementing monocular depth estimation [114].	50
Figure 3.1 Flowchart depicting the core structure of the NS-SFR framework, which extracts, isolates, and validates step-edges from natural scene images before measuring the NS-SFR from a natural scene image dataset. [124].	57
Figure 3.2 The ROI selection process, adjusting the box boundaries for cropping the ROIs, measuring the ESF mask and applying the pixel stretching (cf. § 3.3.2), then validating that the selected ROIs contain the required step-edge profile (cf. § 3.3.2). The position in the framework is shown in Figure 3.1, marked with <i>A</i> . [124].	65
Figure 3.3 Depiction of the ESF mask [76].	66
Figure 3.4 Flowchart highlighting the steps for the pixel stretching filter. The position in the framework is shown in Figure 3.2, marked with <i>A.1</i> . [124].	67
Figure 3.5 Illustration of the <i>T</i> -shaped median coordinates and how they were used to fill each pixel row, i.e. stretching the median pixel value. [76].	68
Figure 3.6 An example of edge isolation using the pixel stretching on simulated ROIs, one noiseless and the other with a noise level of SNR 4. [76].	69
Figure 3.7 These two graphs illustrate the effect that the pixel stretching edge isolation has on ROI with high noise, SNRs 25, 12, 6 and 3. On the left are the resultant e-SFRs from these ROIs without pixel stretching, and on the right are e-SFRs from the same ROIs but applying pixel stretching [124].	70
Figure 3.8 A demonstration of how the pixel stretching edge isolation reduces the effects of non-uniform illumination in 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 [124].	71
Figure 3.9 Three examples of false sharpening occurring in natural scene ROIs. Each example is illustrated without and with pixel stretching, and the pixel stretched ESFs and NS-SFRs are plotted.	72
Figure 3.10 These seven ROIs contain various edge profiles, including step edges, staircase edge, line edge, roof edge, and trough edge. This diagram describes how the measured gradient can successfully determine useable step edge ROIs [76, 124].	73
Figure 3.11 Step-edge validation flowchart describing the process to select only step-edge ROIs. This process is marked with <i>A.2</i> in the framework in Figure 3.2 [124].	74
Figure 3.12 Depiction of the step-edge verification process. 0.1 is the gradient threshold in this example, any gradient above this threshold is marked.	75
Figure 3.13 Simulated demonstration of the two vertical gradients following the ESF mask boundary. These gradients are used to detect areas of non-uniformity in the isolated ROIs.	76
Figure 3.14 Simulated illustration of the detection and segmentation of unwanted changes in the a) <i>x</i> and b) <i>y</i> edge directions.	77

Figure 3.15	Flowchart depicting measuring the NS-SFR from the validated ROIs. This process is marked with 'B' in the framework in Figure 3.1 [124].	78
Figure 3.16	The e-SFR envelope compares the manual isolation of test chart edges with those extracted using the NS-SFR framework [76].	80
Figure 3.17	The NS-SFR envelope from four natural scenes compared to the equivalent test chart e-SFR envelope. The envelopes are outlined by the 5 th and 95 th percentiles of the data and a weighted mean [76].	82
Figure 3.18	The mean of the weighted average NS-SFR envelopes of 30 natural scenes compared to the mean of a test chart e-SFR envelope [76].	83
Figure 4.1	Examples of simulated square ROIs (128x128 pixels) ranging in angle and contrast. These examples are noiseless and without pixel stretching.	87
Figure 4.2	Examples of simulated ROIs with the height of 128 pixels ranging in contrast and ROI width. These examples are noiseless and without pixel stretching.	88
Figure 4.3	Examples of simulated ROIs with the width of 128 pixels ranging in contrast and ROI height. These examples are noiseless and without pixel stretching.	88
Figure 4.4	Examples of a captured step-edge test chart, at Michelson contrasts of 0.12, 0.52 and 0.92, positioned at 2.5°, 25° and 42.5°.	89
Figure 4.5	The relationship between the number of electrons and the pixel value. The trend is linear for a gain value of 1 (no gain), reaching the FWC. When the gain is applied, the well capacity decreases, requiring fewer electrons to reach pixel saturation. Adapted from [139].	91
Figure 4.6	A simulated 5° step-edge in a 128x128 pixel ROI window, with the six levels of image noise emulation used in this study.	93
Figure 4.7	Comparison of the simulated MAE introduced by the SFR parameter ranges, edge angle, edge contrast, ROI height and ROI width, between spatial frequencies 0.37 to 0.50 cyc/pixel, with and without applying pixel stretching at seven noise levels.	95
Figure 4.8	Matrix of simulated noiseless variation introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by the individual parameter ranges.	96
Figure 4.9	Matrix of simulated SNR 18 variation introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by the individual parameter ranges [145].	96
Figure 4.10	Comparison of the DSLR 1 captured test chart MAE introduced by the SFR parameter ranges, edge angle, edge contrast, ROI height and ROI width, between spatial frequencies 0.37 to 0.50 cyc/pixel, with and without applying pixel stretching at six SNRs.	98
Figure 4.11	Matrix of variance introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by each SFR parameter. The data is from test chart ROIs, captured using the DSLR1 at a gain setting of ISO100 (SNR 97).	99

Figure 4.12 The optimal edge angles (a) 0° and b) 45°) for recording the modulation with a sensor array, demonstrated with a sine-wave test chart with the frequency of 0.27cyc/mm. Adopted from [20].....99

Figure 5.1 Flowchart depicting the proposed method to estimate the system e-SFR from natural scene derived SFR data.105

Figure 5.2 The horizontal and vertical angles were converted for each of the 16 diagonal frame subdivisions used to obtain the sagittal and tangential edges.106

Figure 5.3 These histograms illustrate LSF FWHM from the test chart and NS-SFR step-edges for data from the DSLR 1 TIFF image dataset. The histograms are normalised using their peak value. This example illustrates distances 1, 2 and 6 out of 6 radial segments. The red bars indicate the 10th, 20th and 50th percentiles for the NS-SFR LSF FWHM distribution [124].107

Figure 5.4 The comparison of MAE introduced by e-SFR parameters, edge angle, edge contrast and ROI height, between spatial frequencies 0.37 to 0.50 cyc/pixel, from NS-SFRs calculated using three LSF FWHM thresholds for three systems [124].112

Figure 5.5 Matrix of variation introduced to the NS-SFR measurement when changing edge angle, edge contrast, ROI height and ROI width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots MAE introduced by each SFR parameter. Data is from the DSLR 1 system using a TIFF pipeline.112

Figure 5.6 The distribution of the ISO gain settings for the images in the three datasets [146].114

Figure 5.7 Scene type distributions for the three datasets, classified by re-trained AlexNet. Adopted from [146].115

Figure 5.8 Number of ROIs isolated and the number of utilised ROIs per radial distance for the vertical e-SFR estimation from TIFF DSLR 1 dataset for each LSF FWHM distribution threshold.117

Figure 5.9 Number of ROIs isolated and the number of utilised ROIs per radial distance for the sagittal e-SFR estimation from TIFF DSLR 1 dataset for each LSF FWHM distribution threshold.118

Figure 5.10 Absolute error at the 25%, 50% and 75% of the system vertical e-SFR estimate Nyquist frequency, derived from three LSF FWHM thresholds, for the DSLR systems at radial segments 1/6, 3/6 and 6/6 as well as the weighted average (WAve.) of the entire frame. Adopted and expanded on from [146].120

Figure 5.11 DSLR 1 vertical system e-SFR estimation for three radial distances out of six and a weighted mean of all six radial distances. 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 mean of the ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [145, 146].122

Figure 5.12 DSLR 2 vertical system e-SFR estimation for three radial distances out of six and a weighted mean of all six radial distances. 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 mean ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [146].123

Figure 5.13 Comparison between the DSLR 1 (TIFF and RAW datasets) vertical weighted average system e-SFR estimates derived from three LSF FWHM distribution thresholds, 10 th , 20 th and 50 th percentiles.	125
Figure 5.14 Comparison between the DSLR 2 vertical weighted average system e-SFR estimates derived from three LSF FWHM distribution thresholds, 10 th , 20 th and 50 th percentiles.	126
Figure 5.15 Weighted average vertical system e-SFR estimate from the DSLR 1 sub-datasets of indoor, man-made and nature scene classes. These estimates are compared to the average ISO12233 slanted edge e-SFR and its standard deviation envelope.....	127
Figure 5.16 25%, 50% and 75% Nyquist frequency of DSLR 1 sagittal TIFF and RAW system e-SFR estimates are plotted against the radial distance. Estimates are derived from the 10 th , 20 th and 50 th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO12233 e-SFR is also plotted, with a standard deviation of the ISO12233 shaded.	129
Figure 5.17 25%, 50% and 75% Nyquist frequency of DSLR 2 system e-SFR estimates are plotted against the radial distance. Estimates are derived from the 10 th , 20 th and 50 th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO12233 e-SFR is also plotted, with a standard deviation of the ISO12233 shaded.....	130
Figure 5.18 Absolute error at the 25%, 50% and 75% of the smartphone vertical e-SFR estimate Nyquist frequency, derived from three LSF FWHM thresholds, at radial segments 1/6, 3/6 and 6/6 as well as the weighted average (WAVE.) of the entire frame. Adopted and expanded on from [146]......	131
Figure 5.19 The Smartphone vertical system e-SFR estimation for three radial annuli out of six and a weighted mean of the entire field of view. 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 mean ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [146]......	132
Figure 5.20 Weighted average vertical system e-SFR estimate from the entire Smartphone dataset. These estimates are compared to the ISO12233 slanted edge method and Spilled-Coins Texture-MTF [145, 146]......	133
Figure 5.21 Weighted average vertical system e-SFR estimate from the Smartphone sub-datasets of indoor, man-made and nature scene classes. These estimates are compared to the ISO12233 slanted edge method standard deviation envelope [146]......	134
Figure 6.1 Methodology used to evaluate the accuracy in system e-SFR estimation when changing input image dataset size and scene content.	138
Figure 6.2 The distribution of the scene locations and lighting conditions of the indoor, man-made and nature datasets used in the AlexNet transfer learning process.	141
Figure 6.3 This figure contains five examples of where the retrained AlexNet predicted the incorrect scene location category.	142
Figure 6.4 Plots depicting the precision and accuracy of the 50% Nyquist system e-SFR estimation, at radial distances 1/6, 3/6, 6/6 and the weighted average of the frame, using various sized subsets. This data is presented for the mixed scene subsets and the three scene classes, man-made (exterior), indoor and nature [146]......	145

Figure 6.5 Plots illustrating the missing data and standard deviation across the seven DSLR 1 subdivisions against the radial distance segments for the mixed scene system e-SFR estimates.....146

Figure 6.6 Plots illustrating the missing data and standard deviation across the seven DSLR 1 subdivisions against the radial distance segments per separated scene class [146].....146

Figure 7.1 The peak absolute error from the slanted-edge (as presented in this thesis) and PSF-CNN (as published by Bauer et al. [15]) approaches, at 10, 20, 30 and 40 lp/mm across the entire frame. This comparison is between similar, but not identical, systems.164

Figure 7.2 The absolute error from the slanted-edge (as presented in this thesis) and PSF-CNN (as published by Bauer et al. [15]) approaches, at 10, 20, 30 and 40 lp/mm for radial distances 1/6, 3/6 and 6/6. This comparison is between similar, but not identical, systems.165

List of Tables

Table 2.1 ROI and edge parameter ranges recommended for the ISO12233 slanted-edge method.....	29
Table 4.1 ROI and edge parameter ranges for the e-SFR ISO 12233 method against those used to estimate the system-SFR from the NS-SFR data.....	100
Table 5.1 Number of test chart edges per radial distance segment for each of the system e-SFR orientations (Vertical (V), Horizontal (H), Sagittal (S) and Tangential (T)).	116
Table 5.2 Percentage utilisation for each dataset used to estimate the vertical system e-SFR for radial distance segments representing the centre, partway, corners, and entire frame.	117
Table 5.3 Percentage utilisation for each dataset used to estimate the sagittal system e-SFR for radial distance segments representing the centre, partway, corners, and entire frame.	119
Table 6.1 Seven subdivisions of the scene classed DSLR 1 dataset [146].	143

Acknowledgements

The journey that has led to the completion of this thesis is a result of efforts and support from numerous people.

I am deeply grateful to my Director of Studies and mentor, Dr Sophie Triantaphillidou, who provided the opportunity for me to pursue this PhD, offering constant guidance and support in my development as a researcher and my education since I was introduced to imaging science in my undergraduate years. I would like to extend my gratitude to the other two supervisory team members, Dr Robin Jenkin and Dr Alexandra Psarrou, for sharing their invaluable knowledge and understanding. The careful comments and constructive criticism given by this excellent supervisory team has shaped this project, allowing me to not only complete this thesis, but also write five published research papers. Thank you for your kind support and encouragement.

A special thanks should be given to Dr Jae Young Park for sponsoring this work financially, making it possible for me to pursue a PhD.

I would not have had the opportunity to complete this project without several mentors along the way. Nour Eddine El Ghoumari introduced me to photography, sparking my passion for imaging. Dr Elizabeth Allen, John Smith, and Dr Efthemia Bilissi taught me the science behind the camera, leading to my enthusiasm for both the art and science of photography.

I must also thank my fellow researchers in the computational vision and imaging technology research group. Including Dr Edward Fry for the engaging conversations and guidance, and Janko Smejkal for the collaboration in developing the image noise models and testing various natural network applications. Also, I am grateful to my fellow PhD researchers in computer science, Mahmoud Aldrainli and Deepika, for their friendship throughout this PhD.

I am deeply thankful for Lorenzo Coccia and Miriam Fox, my closest friends who have helped me through this journey (*Elen sila lumenn' omentielvo*).

Finally, but by no means least, I would like to thank my family, particularly my parents, Dom and Sue van Zwanenberg, for their unfailing support in all I endeavour to achieve.

Author's Declaration

I, the undersigned, hereby declare that this thesis and the work presented in it are my own and result from my original research. Any other sources of information and visual material that have been used, directly or indirectly, have been referenced appropriately.

Oliver van Zwanenberg. B.Sc. (Hons)

January 2022

Chapter 1 Introduction

In this chapter, the Modulation Transfer Function (MTF) is introduced alongside the challenges with the current laboratory approaches. The aims and objectives of this project are detailed, along with the thesis structure and the contributions to knowledge.

1.1 Camera System Performance

Evaluation of camera performance, which determines how accurately the image represents the scene, is essential in various aspects of imaging science. System performance can refer to several areas of image quality, including colour accuracy [1], tone reproduction [2], assessment of geometric distortions [3] and image noise evaluation [4]. The research conducted in this thesis focuses on the MTF [5], which derives a system's resolution and sharpness performance. Resolution describes the level of detail a system can render, and sharpness is the clarity or contrast of the rendered detail, particularly the representation of edges [6, p. 443]. The MTF, resolution and sharpness are not interchangeable; they relate to the micro-image (spatial) properties of a system [6, p. 443]. In this thesis, when referring to performance, it is in terms of the MTF.

The MTF is a widely used evaluation measurement for system performance. It describes the modulation reduction from a camera system as a function of spatial frequency [7, p. 128]. This function forms half of a bell-shaped curve, as illustrated in Figure 1.1; as spatial frequencies increase, the system's ability to render the modulation decreases. It is the rate of change that describes the performance of the system. A high-performance MTF maintains

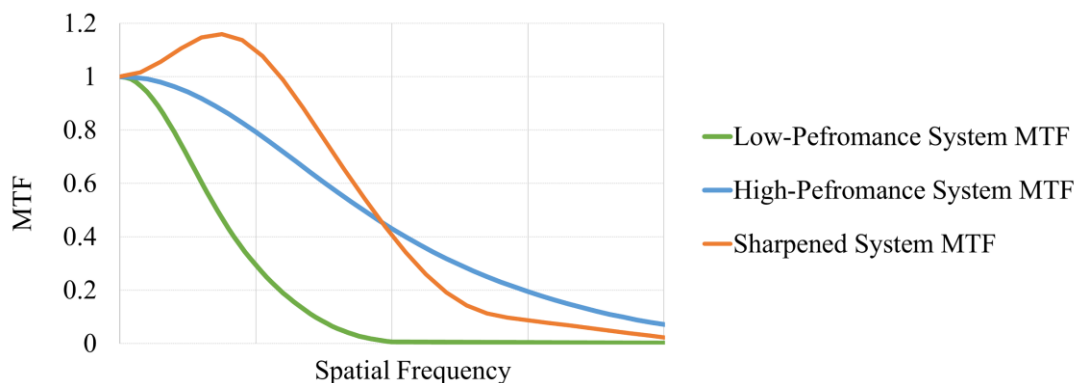


Figure 1.1 Examples of a low-performance MTF, a high-performance MTF and a MTF from a sharpened edge.

higher modulation in the high spatial frequencies than a low-performance MTF. When a sharpening spatial filter is applied to a captured edge, the resultant MTF performance is increased, a low-frequency *lobe* above 1.0 modulation is a sign of sharpening being applied.

Figure 1.2 provides a simple pipeline of a digital camera system. The optical system and sensor are significant contributors to performance and can be considered linear system components. Further, the low pass filter (if present), infrared (IR) cut filter, colour filter array (CFA), micro-lens array, as well as the electronics and signal processing will all affect the output image. The interaction of all components of a system contributes to its performance capabilities. As a result, the system MTF follows the cascading principle [6, p. 449], expressed as:

$$MTF_{sys}(\omega) = MTF_{Lens}(\omega) \cdot MTF_{Array}(\omega) \cdot MTF_{Elec}(\omega) \cdot \dots \quad (1.1)$$

where the system MTF, MTF_{sys} , is the product of the MTF of each component.

Image signal processing (ISP) has become a significant part of a system's performance and, in turn, the visual (output) image quality (IQ). The ISP can compensate for hardware with

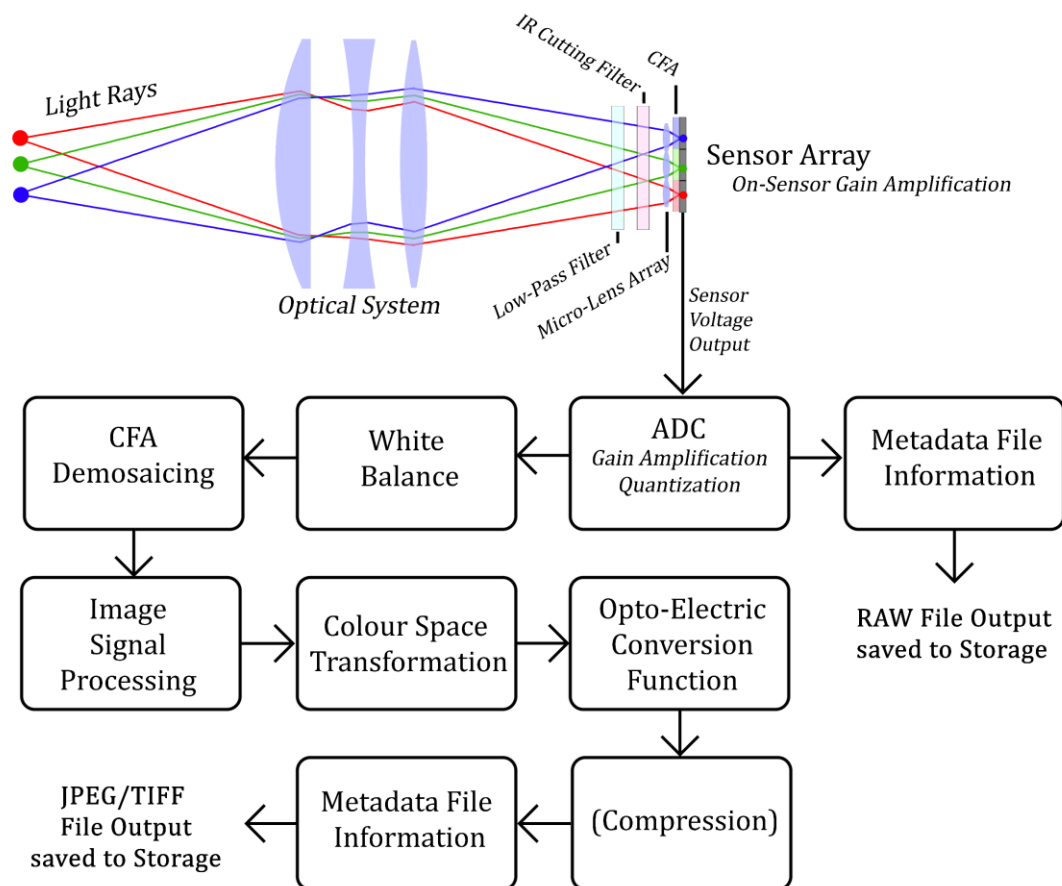


Figure 1.2 Visualisation of a simplified digital camera system pipeline.

significant size and material constraints, such as smartphone camera systems. Examples of the ISP include deep learning IQ enhancements [8], locally adaptive sharpening, denoising and tone mapping, as well as the fusion of multi-image scene capture to obtain higher pixel resolution (digital zoom via super-sampling), noise reduction and improved dynamic range [9]. Modern systems also use subject detection and isolation tools to blur backgrounds, synthetically adding bokeh (a visually pleasing out-of-focus region of a photographic image caused by shallow depth of field, which cannot be obtained with a small lens and sensor design) [9]. All these processes are highly non-linear.

There are several established methods to obtain the MTF, including:

- The sinewave method (cf. § 2.1.2)
- The edge method (cf. § 2.1.3)
- The texture method (cf. § 2.1.4)

These methods are based on linear system theory [10, p. 233], hence, employed under the assumption that the system is linear, homogeneous, and spatially invariant (cf. § 2.1.1).

Despite camera system non-linearities, the MTF is incorporated in camera performance standards [5, 11]. This project has adapted the edge method to obtain natural scene derived performance measurements. The edge method works using the idea that a perfect step-edge contains constant modulation across all spatial frequencies (cf. § 2.1.3). After such an edge is photographed, the system degrades the edge. Taking a trace of the imaged edge provides the system's Edge Spread Function (ESF), $esf(x)$. The differentiation of the ESF provides the system's Line Spread Function (LSF), $lsf(x)$, Equation 1.2. The MTF is calculated through the modulus of the Fourier transform of the LSF, Equation 1.3. [10, pp. 244–246].

$$lsf(x) = \frac{d [esf(x)]}{dx} \quad (1.2)$$

$$M_{edge}(\omega) = \left| \int_{-\infty}^{+\infty} lsf(x) e^{-2\pi i \omega x} dx \right| \quad (1.3)$$

Specifically, the ISO12233:2017 standard slanted-edge algorithm [5] (cf. § 2.3) is utilised in the methodology, an extension of the edge method that returns an edge-based Spatial Frequency Response (e-SFR), which is less susceptible to aliasing and is widely used in the industry.

1.2 The Limitations of Current MTF Approaches

All established MTF methods use a known input signal, which is then compared to the output to provide a measure of signal degradation. Problems occur when non-linear ISP is in the camera system pipeline, resulting in a system with no simple input-to-output relationship. As a result, the system performance will vary depending on what is being photographed and the illumination conditions of the scene, i.e. it becomes scene and processing dependant. Consequently, for the same system, the MTF varies depending on what method is being implemented.

When using a natural scene as an input, there is potential to obtain a scene-and-process dependant MTF (SPD-MTF), more faithfully representing the entire system pipeline (including the ISP). In contrast to purpose-made test charts, natural scene inputs are uncharacterised. Previous works have used natural scenes as test charts, either physically printed [12] or simulated [13, 14] (cf. § 2.8.2), establishing the required input characteristics. However, such methods do not provide the possibility for a *no-reference* real-world implementation, i.e. a performance measurement with an unknown input. Alternatively, the MTF can be estimated by training a neural network on characterised scenes [15, 16]; such approaches have several benefits but also have various drawbacks (cf. § 2.8.4).

If a *no-reference* performance measurement is established, it opens up new possibilities in imaging science, moving camera characterisation away from laboratory and test chart approaches. There are many possible applications; for example, Deep Neural Networks (DNN) are currently used as one of the main technologies in image recognition tasks. With the application of DNN in decision critical systems, such as autonomous vision systems, the development of camera performance measures derived from natural scenes would allow the camera output quality to be monitored at any given moment, i.e. in real-time.

1.3 Scope of Work

The project investigates and develops a novel methodology that estimates the MTF/SFR directly from pictorial natural scenes rather than the standard test charts. The research presented in this thesis aims to take the slanted-edge algorithm away from laboratory conditions and apply it to natural scene inputs to obtain an e-SFR estimation comparable to

the ISO standard method. Natural scene derived camera system e-SFR measurements open possibilities to real-time SFR monitoring, suitable for applications related to intelligent vision systems, as well as provide a potential for scene dependant performance analysis.

The slanted-edge method was chosen for this task over alternatives as, firstly, edges are readily available in most natural scenes. Secondly, the input scene does not need to be characterised for the algorithm to operate (as long as it is of the correct edge profile), meaning *no-reference* performance measurements are potentially possible. Finally, if successful, the natural scene derived SFR can be directly compared to the ISO12233 e-SFR.

The following research objectives are achieved in this thesis:

- 1) Develop an automated methodology to locate, isolate and verify step-edges from natural scenes. These selected edges pass through the slanted-edge algorithm to produce natural scene derived SFR (NS-SFR).
- 2) Use NS-SFRs to estimate the camera system e-SFR, measured according to the ISO12233 standard.
- 3) Determine the best-suited edge parameter ranges that result in a stable system e-SFR estimation.
- 4) Implement and evaluate this novel approach and provide expected accuracy and precision for linear and non-linear ISP cameras. Investigate the type and number of images (dataset sizes) for the accurate estimation of the e-SFR.

Potential applications would include:

- Camera system comparisons without the capture system at hand, utilising online datasets.
- The development of SPD-MTFs that can be used to improve the performance of spatial image quality metric (IQM) that incorporate the MTF/SFR [13, 14, 17].
- Monitoring the performance of deployed imaging vision systems over time, such as autonomous vehicles and live security systems.

1.4 Thesis Structure

The structure of this thesis is as follows:

Chapter 2 introduces camera MTF/SFR measuring techniques and associated test charts, digital image formation, the slanted-edge method, and sources of variation and error. Following this, how SFRs/MTFs are analysed, recent research in improving the ISO12233 standard, current solutions to obtaining system performance from natural scenes, and finally, image noise analysis are all detailed. This critical literature review provides background information and current research that influenced the development of the proposed methodology presented in the subsequent chapters.

A framework to extract, isolate and verify natural scene step-edges is presented in Chapter 3. This chapter includes studies carried out to identify the most suitable edge detection algorithm, determine how ROIs are cropped using adaptive window sizes, and establish edge isolation techniques. Further, it details the process of using the extracted step-edges in the ISO12233 slanted-edge algorithm. The chapter concludes with initial NS-SFR results and observations.

Natural scenes contain edges of various edge angles, contrasts and region of interest (ROI) window sizes required for isolation. These numerous parameters and their combinations introduce large variations in the NS-SFR data. Chapter 4 details research into the ISO12233 e-SFR input parameter ranges, expanding upon previous publications. The aim is to provide a range of edge and ROI characteristics that return stable system e-SFR estimations without excessively restricting NS-SFR data, from which the suitable for purpose step-edges are rare.

The methodology to estimate the system e-SFR from NS-SFRs is detailed in Chapter 5. This methodology is applied to three diverse camera image datasets, assessing the accuracy of the method compared to the equivalent ISO12233 e-SFR. The first two camera systems are close-to-linear, one equipped with a wide-angle and the other a telephoto lens. The third is a smartphone camera system with a highly non-linear ISP. Results and observations concerning each dataset are discussed.

Chapter 6 evaluates the accuracy and precision of the system e-SFR estimation across the camera frame when using different dataset sizes, with the number of images ranging from

10 to 570. In addition, the datasets are grouped by the scene location. This study evaluates the number of images and the scene content suitable for the proposed methodology.

Detailed discussions on the advantages and caveats of the proposed methodology, the implementation, recommended thresholds, comparison to alternative approaches and potential applications are provided in Chapter 7.

Chapter 8 states the conclusions of the thesis and recommends further work.

1.5 Contributions to Knowledge

Research in this thesis results in the following original contributions to knowledge:

- A novel automated framework that extracts step-edges from natural scenes that are suitable for implementation with the slanted-edge algorithm. It includes the creation of a novel isolation technique that, not only improves the yield of natural scene edges, but also reduces the effects of noise and non-uniformity on the resulting NS-SFR.
- Suitable edge and ROI parameter ranges are established to obtain a stable system e-SFR estimations without excessively restricting NS-SFR data. These ranges are appropriate for edges from test charts as well as natural scenes.
- The development of a novel methodology to estimate the system e-SFR from NS-SFRs, with accuracy comparable to the ISO12233 e-SFR.
- An understanding of the expected accuracy and precision of system e-SFRs estimated from natural scenes for various camera systems and scene types.

1.6 Related Work

1.6.1 Publications

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, pp. 60402-1-60402-16, Vol. 65 Issue 6, 2022

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, pp. 60405-1-60405-12, Vol. 65 Issue 6, 2022

O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, ‘*Natural Scene Derived Camera Edge Spatial Frequency Response for Autonomous Vision Systems*’ IS&T/IoP London Imaging Meeting, pp. 88-92, 2021.

O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, ‘*Camera System Performance Derived from Natural Scenes*’ IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XVII, pp. 241-1-241–10, 2020.

O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, ‘*Edge Detection Techniques for Quantifying Spatial Imaging System Performance and Image Quality*’, IEEE: New Trends in Image Restoration and Enhancement (NTIRE) workshop, in conjunction with Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1871-1879, 2019.

1.6.2 Presentations at Conferences and Symposia

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Estimation of ISO12233 edge spatial frequency response from natural scene derived step-edge data*’, at IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XIX, 2022, San Francisco, USA (Online).

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Analysis of natural scene derived spatial frequency responses for estimating the camera ISO12233 slanted-edge performance*’, at *IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XIX*, 2022, San Francisco, USA (Online).

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Camera Performance Derived from Natural Scenes*’ at IS&T Seminar Series - Best Student Research 2020 (Online).

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Measuring Camera Performance from Natural Scenes*’ at the University of Westminster, School of Computer Science and Engineering Doctoral Conference, 2020, (Online Poster Presentation).

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Measuring Camera Performance from Natural Scenes*’ at *IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XVII*, 2020, San Francisco, USA.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Measuring Camera Performance from Natural Scenes*’ at The Royal Photographic Society Imaging Science Group Symposium, Good Picture 'Imaging Revealed', 2019, University of Westminster, Regent Street Campus, London, UK.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Image System Performance measure through Edge Extraction*’ at the University of Westminster, School of Computer Science & Engineering Student-Staff Research Conference, 2019, New Cavendish Street Campus, London, UK.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Edge Detection Techniques for Quantifying Spatial Imaging System Performance and Image Quality*’ at *New Trends in Image Restoration and Enhancement (NTIRE)* workshop Poster Presentation, in conjunction with *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, Long Beach, California, USA.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Quantifying Camera Performance from Natural Scenes using Edge Detection Techniques*’ at the University of Westminster, School of Computer Science and Engineering Doctoral Conference, 2019, New Cavendish Street Campus, London, UK.

S. Triantaphillidou (*presenter*), E. W. S. Fry (*presenter*), and O. van Zwanenberg (*presenter*), “*Objective Image Quality Assessment Using Pictures,*” at Transactions Imaging/Art/Science: Image Quality, Content and Aesthetics, 2019, University of Westminster, Regent Street Campus, London, UK.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Camera System Performance Measure through Edge Extraction*’ at the University of Westminster Home Office Poster Presentations, 2019, Regent Street Campus, London, UK.

O. van Zwanenberg (*presenter*), S. Triantaphillidou, R. Jenkin, and A. Psarrou, ‘*Modulation Transfer Function Determination from Pictorial Natural Scenes*’ at the University of Westminster, Faculty of Science and Technology Doctoral Conference, New Cavendish Street Campus, London, UK.

1.6.3 Awards

Best Paper Award for the Image Quality and System Performance (IQSP) *XVII* conference, 2020, at the Society for Imaging Science and Technology’s (IS&T) Electronic Imaging Symposium, 2020, San Francisco, USA. Awarded for the paper entitled “*Camera System Performance Derived from Natural Scenes*”.

Chapter 2 Background to System Performance

In this chapter, the theory needed to evaluate camera system performance is discussed, laying the foundation for the subsequent chapters. The topics include measurement of the MTF/SFR, image formation, the slanted-edge method for SFR evaluation, and sources of measurement error. Previous approaches for deriving system performance from natural scenes are discussed, and the literature is reviewed with the premise of developing a natural scene derived Spatial Frequency Response (e-SFR) measure.

2.1 The Modulation Transfer Function

The MTF describes the modulation reduction for a camera system as a function of spatial frequency [7, p. 128]. This function forms half of a bell-shaped curve, as illustrated in Figure 1.1; as spatial frequencies increase, the system's ability to render the modulation decreases. It is the rate of change that describes the performance of the system. A high-performance MTF maintains higher modulation in the high spatial frequencies than a low-performance MTF. This section introduces the fundamentals of MTF measurement. It covers the sine-wave, edge and texture-MTF methods. Figure 2.1 depicts four common test chart patterns used to measure MTF.

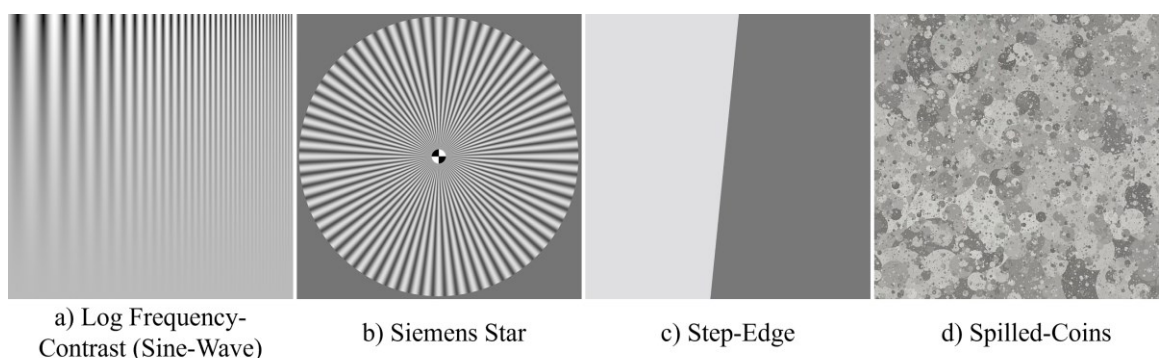


Figure 2.1 Test chart input patterns for MTF and SFR measurements.

2.1.1 System Requirements

MTF is measured through various well-established methodologies, that are traditionally obtained from captured test charts under laboratory conditions. MTF measurement is founded on linear system theory [10, p. 233]. Thus, it is employed under the assumption that the system is linear, homogeneous, and spatially invariant.

A linear system has a simple input-output relationship [18, p. 206]. If two inputs, f and g , are input to a linear system, H , the output would be the same as if the inputs were added before passing the system, expressed as [18, p. 205, 19, p. 503]:

$$H[f + g] = H[f] + H[g] \quad (2.1)$$

A homogeneous system has a linear transformation mapping one plane (the two-dimensional (2D) signal) to another (the sensor array). In other words, when the input, f , is multiplied by a scalar value, a , the output will be the same as if the system output was multiplied by a . This is expressed as [19, p. 503]:

$$H[af] = aH[f] \quad (2.2)$$

The inputs f and g in these examples can either be individual pixels or 2D image functions.

A spatially invariant system, also known as stationary, is where a shift in the input signal will result in an equivalent shift in the output image [7, p. 124].

In practice, however, most camera systems violate these requirements. Many systems utilise non-linear ISP and demosaicing algorithms. The quantisation of a sensor array makes the camera system both non-linear and non-stationary. Also, optics that contain any aberrations are non-stationary. That said, imaging systems are commonly treated as *quasi*-linear and spatially invariant under controlled conditions and restricted sensor areas [7, p. 124].

2.1.2 The Sinewave Method

The first MTF measuring technique is the sinewave method, one of the earliest measurements of MTF. In its simplest form, it is derived from patches of sine wave patterns of various known frequencies. The ratio between the modulation of the output and input provides the MTF [10, pp. 241–244], as expressed in Equations 2.3 and 2.4 [20].

$$m(\omega) = \frac{Max - Min}{Max + Min} \quad (2.3)$$

$$M_{sine}(\omega) = \frac{m(\omega)_{OUT}}{m(\omega)_{IN}} \quad (2.4)$$

where $m(\omega)$ is the modulation of a sinusoidal wave with frequency ω , Max and Min are the maxima and minima of the sinusoid, and $M_{sine}(\omega)$ is the sinewave MTF, calculated by the input and output modulation.

More complex extensions of the sinewave method include the frequency sweep and the Siemens star methods [5, 21].

Figure 2.1 a) illustrates the Imatest Log Frequency test chart [21] used in the frequency sweep method. The spatial frequencies increase on a logarithmic scale in the chart, allowing the sinewave approach to be implemented across known changing frequencies. This chart also incorporates decreasing contrast along the y axis to measure system aliasing (cf. § 2.2.2).

Figure 2.1 b) illustrates the Siemens star test chart. The ISO12233:2017 [5] standardises the Siemens star approach and refers to it as the sine-based Spatial Frequency Response (s-SFR).

2.1.3 The Edge Method

Like the sinewave method, the edge method originated from analogue film photography. It works from the premise that a perfect step-edge contains infinite spatial frequencies [6, p. 446], as demonstrated in Figure 2.2. When such an edge is photographed, the system degrades it. Taking a trace of the imaged edge provides the system's Edge Spread Function (ESF), $esf(x)$. The differentiation of the ESF provides the system's Line Spread Function (LSF), $lsf(x)$, Equation 1.2. The MTF is calculated through the modulus of the Fourier transform of the LSF, Equation 1.3. [10, pp. 244–246].

A MTF must be measured from a perfect step-edge with respect to the system, i.e. a *near*-perfect edge with a modulation equal to one across the camera system spatial frequency

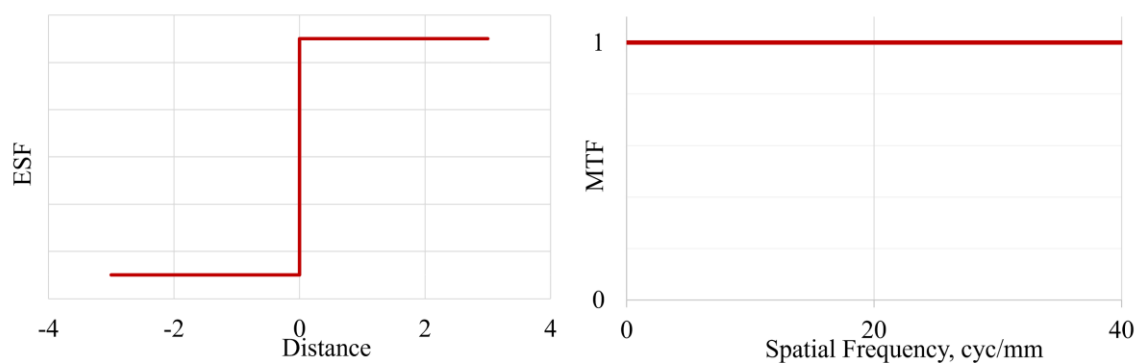


Figure 2.2 Illustration of a perfect step-edge, containing maximum modulation across all spatial frequencies.

bandwidth, as depicted in Figure 2.3. If the input edge is not *near*-perfect, it will not yield a pure MTF. Suppose the step-edge contains higher modulation than the capturing capabilities of the system, as shaded in Figure 2.3, it is suitable to obtain the MTF but must be corrected by dividing the output by the input edge MTF, see Equation 1.1. When the step-edge contains lower modulation than the capturing capabilities of the system, it is unsuitable for a MTF measurement.

With the introduction of digital systems, the edge method became impractical due to the edge having to be perfectly aligned with the pixel array (cf. § 2.2.3) and aliasing becoming problematic. In response, the slanted-edge algorithm was established [5, 22, 23], as detailed later in this chapter (cf. § 2.3), utilising a slanted step-edge test chart, Figure 2.1 c). This method solved many problems with the edge method for digital systems, allowing a simple implementation and in-phase measurements. The output is not strictly speaking an MTF; instead, it is referred to as an edge-based Spatial Frequency Response (e-SFR) [5]. This is due to two reasons. First, the resampling of the ESF (cf. § 2.3.2) obtains a measurement above the Nyquist frequency, removing aliasing issues. Secondly, the modulation content of the input edge is not corrected for in the measurement. The e-SFR is measured with the assumption that the input step-edge is perfect or near-perfect. The step-edge test chart is printed and photographed at a set distance to meet this condition (cf. § 2.3).

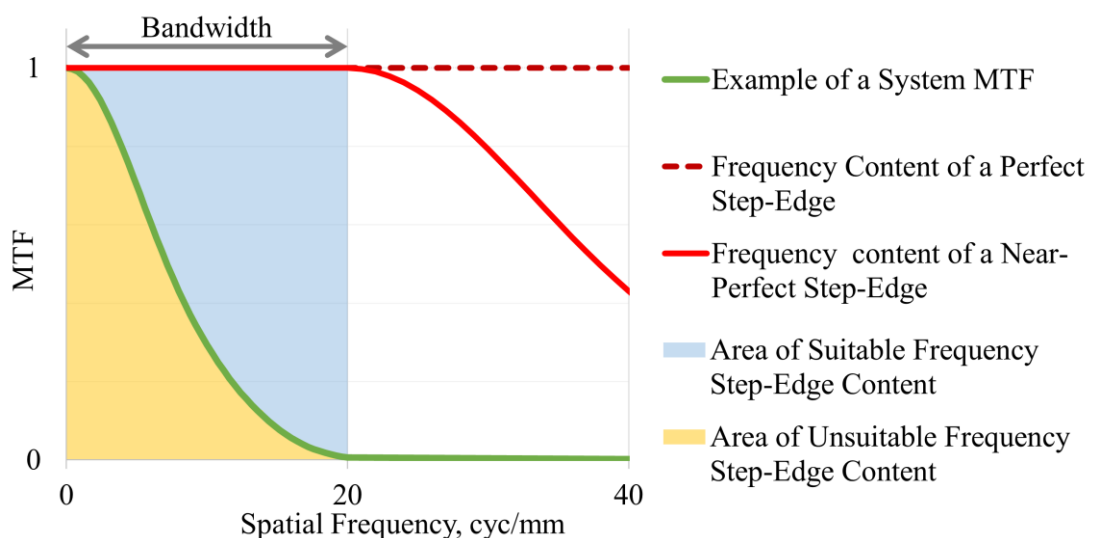


Figure 2.3 MTF plots depicting the frequency content of a perfect step-edge and a near-perfect step-edge, with respect to an example of a system MTF. The shaded area above the example system MTF illustrates the region where the frequency content of an input step-edge is *suitable* for measuring the MTF but requires correction. The shaded area below the system MTF illustrates the frequency content *unsuitable* for measuring the MTF.

2.1.4 The Texture Method

Many modern systems apply non-linear adaptive ISP, which results in scene dependent system performance [14, 20, 24]. For instance, different levels of sharpening and denoising may be added in response to the signal, scene illumination and camera settings. This means the performance changes based upon what is being photographed. Furthermore, the ISP is set by a manufacturer, and for many systems, it comes as an unknown *black box*. As a result, the sinewave and edge methods would produce different responses from such non-linear ISPs, neither representing the output performance from a natural scene. Thus, the texture-MTF method was developed, using a test chart pattern that contains the statistical properties of an average natural scene, [25–29], Dead-Leaves [30, 31] and Imatest’s Spilled Coins chart pattern [32] (Figure 2.1 d)) are commonly used. This method is designed to be scale, shift, exposure and rotation invariant, texture-like, and robust to denoising [25].

The texture-MTF is calculated by taking the square root of the output power spectrum (PS_{out}) over the input (PS_{in}) [25]. The method was further refined by removing the system noise bias by subtracting the output noise power spectrum (NPS_{out}) of the system from the PS_{out} , expressed as [29]:

$$M_{txt}(\omega) = \sqrt{\frac{PS_{out}(\omega) - NPS_{out}(\omega)}{PS_{in}(\omega)}} \quad (2.5)$$

The NPS_{out} is commonly calculated from a captured uniform patch.

More recently, the texture-MTF method is obtained using cross-convolution, first proposed by Kirk et al. [33]. Using the convolution theorem (cf. § 2.2.2), the output, $Y(\omega)$, of a system is given by the product of the Fourier transform of the input, $X(\omega)$, and the transfer function of the system, $H(\omega)$, [33, 34]:

$$Y(\omega) = X(\omega)H(\omega) \quad (2.6)$$

The output power spectrum is calculated from the amplitude response, $|H(\omega)|^2$, and the input power spectrum [33, 34]:

$$PS_{out}(\omega) = E[|H(\omega)|^2|X(\omega)|^2] = |H(\omega)|^2PS_{in}(\omega) \quad (2.7)$$

where $E[]$ is the expected value of the amplitude of the output signal. In turn, Equation 2.7 can be rearranged to form Equation 2.5.

$$M(\omega) = |H(\omega)| = \sqrt{\frac{PS_{out}(\omega)}{PS_{in}(\omega)}} \quad (2.8)$$

The cross-convolution approach aims to provide an MTF that also includes phase information. This is achieved by using the cross-correlation power density, $PD(\omega)$, from the input and output signal [33, 34]:

$$\begin{aligned} PD(\omega) &= Y(\omega) \otimes X(\omega) = E[Y(\omega)X(\omega)^*] = E[Y(\omega)X(\omega)X(\omega)^*] \\ &= H(\omega)E[|X(\omega)|^2] = H(\omega)PS_{in}(\omega) \end{aligned} \quad (2.9)$$

where $*$ denotes the complex conjugate. Thus:

$$H(\omega) = \frac{PD(\omega)}{PS_{in}(\omega)} \quad (2.10)$$

2.1.5 Results and Implementation of the MTF Measurements

All MTF measuring methods should theoretically provide converging performance measurements for an ideal noiseless, linear, and stationary system. In practice, these performance measurements vary [35]. Digital camera systems are non-stationary; the signal is rarely in phase with the sampling grid, and the frequency response degrades when out of phase [20, pp. 82–83, 36, 37]. Furthermore, the source and level of inherent error from each method is different [38, 39, 40, p. 22].

2.2 Digital Image Formation

In this section, the theoretical basics of the formation of an image by camera optics and a sampling array are explained. Then, the process of modelling the expected system MTF is discussed.

2.2.1 Imaging Equation

As light passes through the camera system, it is spread, blurring the image. The spread measure of a single point of light (an impulse) passing through the system is known as the

Point Spread Function (PSF), $I(x, y)$, and it determines the system resolution and sharpness [7, pp. 125–128, 41, p. 56]. The greater the system blurs the impulse, the wider the spread in the x y directions, thus, the lower the system performance.

The LSF is the integral of the PSF in one orientation [7, p. 126]. As the sampling process of a digital system is anisotropic, i.e. it does not have the same physical properties in all directions, its PSF is not rotationally symmetrical. Hence, the LSF depends upon the direction of the integration. Figure 2.4 provides the mathematical relationship between the PSF, LSF, ESF, Optical Transfer Function (OTF) and MTF, adapted from [7, p. 133].

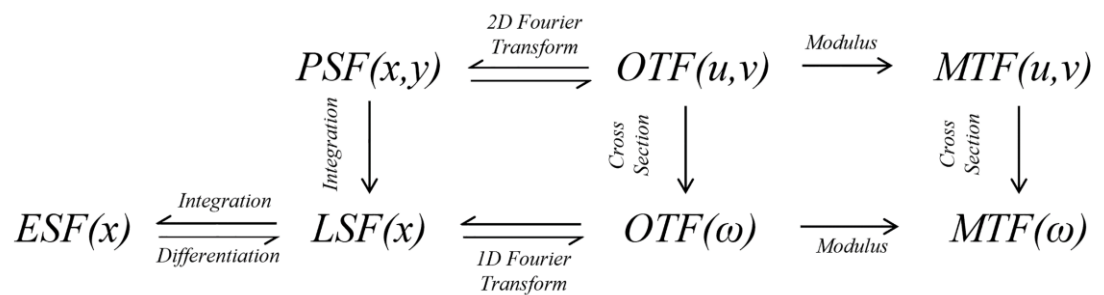


Figure 2.4 The fundamental relationships between spread functions and system performance, adapted from [7, p. 133].

Provided the system is linear and stationary, the image output, $g(x, y)$, can be expressed as the input $f(x, y)$, convolved with the PSF [18, pp. 206–209]:

$$g(x, y) = \int \int_{-\infty}^{+\infty} f(x_1, y_1) I(x - x_1, y - y_1) dx_1 dy_1 \quad (2.11)$$

Which can be expressed in one direction in terms of the LSF as [7, p. 128]:

$$g(x) = \int_{-\infty}^{+\infty} f(x_1) lsf(x - x_1) dx_1 \quad (2.12)$$

2.2.2 Digital Sampling

A Dirac Comb, also known as a Shah function, is a series of impulses, $\delta(x)$, and can represent a one-dimensional (1D) digital array [42]. The separation between each impulse, ΔX , represents the sensor's pixel pitch. This function is described mathematically by Equation 2.13 [7, p. 134].

$$III\left(\frac{x}{\Delta X}\right) \quad (2.13)$$

The discrete sampled function, $\tilde{f}(x)$, is obtained by multiplying the Dirac Comb and the input function. This process is denoted by the following equation [42]:

$$\tilde{f}(x) = f(x)III\left(\frac{x}{\Delta X}\right) = \sum_{n=-\infty}^{\infty} f(x)I(x - n\Delta X) \quad (2.14)$$

A Fourier transformation is a process of converting a real-space function into its sinusoidal components [18, pp. 190–191, 43, pp. 13–14]. The Fourier Transform, $F(\omega)$, is taken of the sampled function using a Discrete Fourier Transform (DFT) [7, p. 130].

The sampled function is substituted with the convolution of the impulse to obtain the DFT of the sampled function [42]:

$$\tilde{F}(\omega) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(x)I(x - n\Delta X)e^{-2\pi i\omega x} dx \quad (2.15)$$

This is simplified by using the convolved sample function, f_n :

$$\tilde{F}(\omega) = \sum_{n=-\infty}^{\infty} f_n e^{-2\pi i\omega \Delta X} \quad (2.16)$$

The sampled function is given by the multiplication of the input signal and the Dirac Comb. Hence, in compliance with the convolution theorem, the function is equal to the DFT of the input signal convoluted by the DFT of Dirac Comb in the frequency domain:

$$F(\omega) \otimes III(u\Delta X) \quad (2.17)$$

In the frequency domain, the impulses of the Dirac Comb are spaced with the interval equal to $\frac{1}{\Delta X}$ [7, p. 134]. When ΔX is too large to sample the signal appropriately, the signal becomes under-sampled. In the frequency domain, the signal becomes overlapped, forming an incorrect sum of the signal, referred to as aliasing. The Nyquist frequency is defined as the spatial frequency equal to half times the inverse of the sampling period [18, p. 197, 44]. This is mathematically described by:

$$\omega_N = \frac{1}{2\Delta x} \quad (2.18)$$

ω_N is the Nyquist frequency, the highest frequency that the system can faithfully reconstruct.

2.2.3 Expected System MTF

A system's MTF can be estimated using calculations that model the expected performance from each camera system component. As the MTF follows the cascading principle, Equation 1.1, each component can be modelled to simulate the MTF of the pipeline [45].

For a diffraction-limited lens, the focal length, f , aperture diameter, D , and light wavelength, λ , determine the resolution capability of the optical system. The focal length and aperture diameter are represented as a f-number ($N = f/D$). A diffraction-limited lens MTF, $MTF_{Lens}(\omega)$, is modelled using [45–47]:

$$MTF_{Lens}(\omega) = \frac{2}{\pi} \left[\cos^{-1} \frac{\omega}{\omega_o} - \frac{\omega}{\omega_o} \sqrt{1 - \left(\frac{\omega}{\omega_o}\right)^2} \right] \quad (2.19)$$

where

$$\omega_o = \frac{1}{\lambda N} \quad (2.20)$$

where ω continues to be the spatial frequency, and ω_o is the frequency of the light.

As the model is based on the input signal wavelength, to obtain a more representative model, the $MTF_{Lens}(\omega)$ can be calculated for a range of wavelengths, weighing the MTF based on the system's quantum efficiency (QE) at each wavelength and then computing an average response [45].

The MTF response from a pixel is a sinc function, based on the 1D pixel size, p , [45]:

$$MTF_{ArrayIN}(\omega) = \frac{\sin(\pi p \omega)}{\pi p \omega} = \text{sinc}(\pi p \omega) \quad (2.21)$$

As mentioned (cf. § 2.1.1), digital camera systems are non-stationary, with performance that varies depending on whether the signal is in or out of phase with the sampling array [20, pp. 76–85, 36, 37]. Equation 2.21 provides the in-phase MTF (maximum system performance). The out-of-phase MTF describes the minimum performance response, calculated by [20, pp. 76–85, 45]:

$$MTF_{ArrayOUT}(\omega) = \left(\frac{\cos(\pi s \omega) \sin(\pi p \omega)}{\pi p \omega} \right) \quad (2.22)$$

where s is the pixel pitch or sampling interval.

The average sensor response is calculated by [20, pp. 76–85, 45]:

$$MTF_{Array_{AVE}}(\omega) = \left(\frac{\cos^2\left(\frac{\pi S\omega}{2}\right) \sin(\pi p\omega)}{\pi p\omega} \right) \quad (2.23)$$

The MTF of the system hardware (optics and sensor), MTF_{sys} , is obtained through multiplying the MTF_{Lens} with the MTF_{Array} , Equation 1.1.

These equations were implemented to simulate the MTF for the *Nikon D800* Digital Single Lens Reflex (DSLR) system with a 24mm lens at an $f/4$ aperture, one of the characterised systems used in the research presented in this thesis (cf. § 5.3.1). This system has a pixel pitch of $4.87\mu\text{m}$. The MTF_{Lens} was calculated using a wavelength of 550nm. The in- and out-of-phase MTF_{Array} were calculated and multiplied with MTF_{Lens} to provide a simulated maximum, minimum and average MTF_{sys} , plotted in Figure 2.5.

The models given above do not account for other camera system components, such as the low-pass filter, micro-lens array, demosaicing, analogue-to-digital conversion (ADC) and ISP, see Figure 1.2, which further degrade the signal. In addition, the simulation does not illustrate performance degradation due to the natural performance variation of the optical system across the field of view (cf. § 2.4.3), crosstalk, image noise, and other optical artefacts, including chromatic aberrations and lens element misalignment. Therefore, the

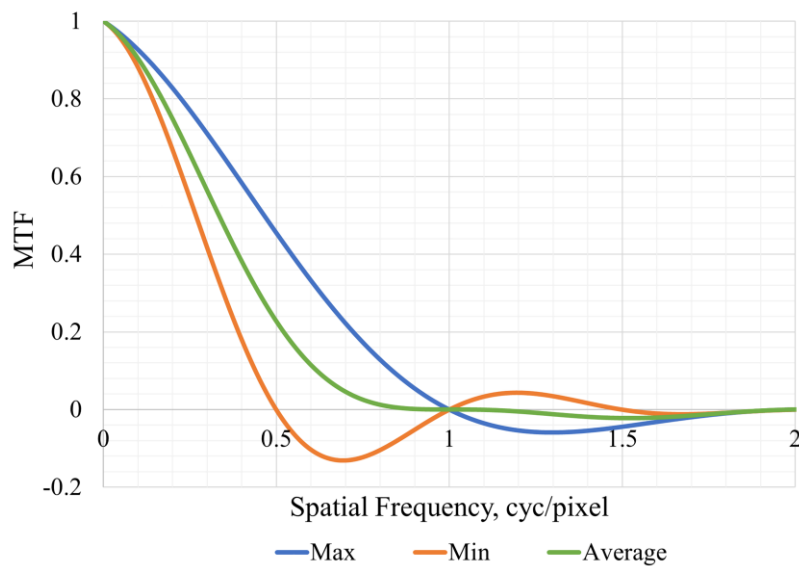


Figure 2.5 Modelled in-phase (max), out-of-phase (min) and average MTF_{sys} for the *Nikon D800* DSLR system up to four times the Nyquist limit.

real-world performance would be lower than these modelled MTFs; if required, more complex models can be incorporated into a system simulation [47, 48]. That said, the above models offer a guideline to engineers that require an ideal system MTF range.

A practical measurement for this project is to convert such simulated MTFs to be in the spatial domain to provide the spread of a given imaged edge, i.e. the ESF or LSF, allowing the minimum distance between neighbouring edges to be determined without their ESFs overlapping (cf. § 3.3.1). This is achieved using the spread function conversions stated in section 2.2.1. Before converting the MTF, it first needs to be mirrored around the zero frequency value and then padded with zeros on either side of the MTF to avoid introducing aliasing and high frequencies [45]. Figure 2.6 provides the maximum and minimum LSFs and ESFs for the modelled system. The resultant edge spread for this system is expected to be between 1.9 and 3.7 pixels, the distance between the points at 50% of the LSF maxima, i.e. the Full Width Half Maxima (FWHM).

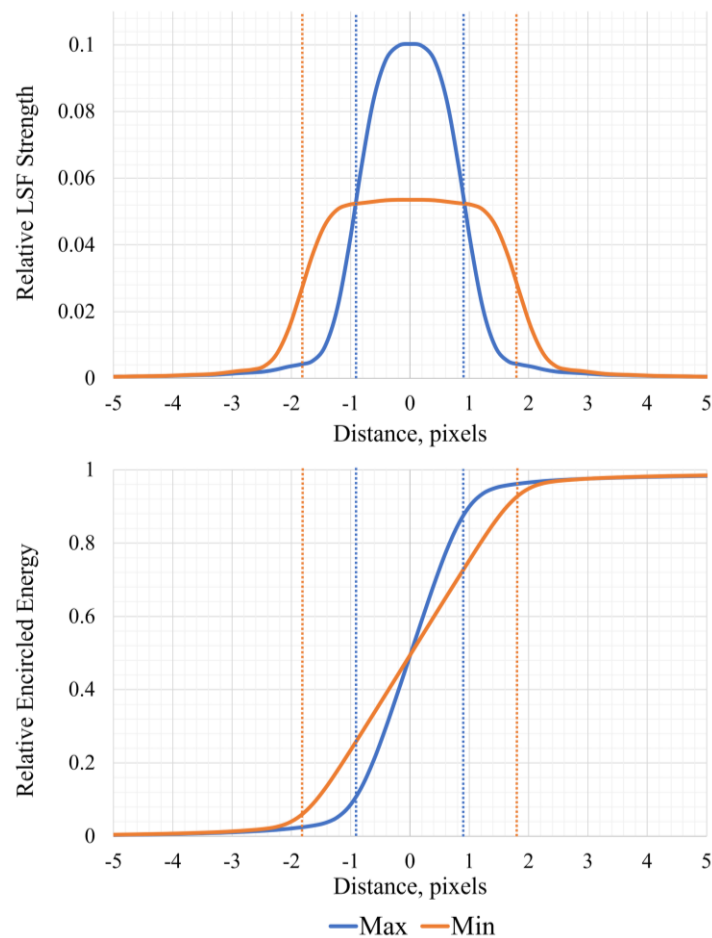


Figure 2.6 The maximum and minimum a) LSF and b) ESF for the simulated *Nikon D800* DSLR system. The spread of the system is marked with vertical broken lines, depicting the point at which the LSFs reaches 50% of the maximum peak (FWHM).

2.3 The Slanted-Edge Method

The slanted-edge method is the standardised procedure to measure the system e-SFR, as detailed in the ISO12233:2017 [5]. Figure 2.7 depicts this method as a flowchart, adapted from [5].

A straight greyscale edge tilted at an angle is photographed under controlled conditions. The illumination across the target should stay within 2% variation [5]. Also, the distance from the target to the sensor should be set to ensure that the input printed step-edge is near-perfect for the e-SFR measurement (cf. § 2.1). This distance depends on the resolution of the printed target and the sensor array. Generally, 1cm of a photographic inkjet print should be equivalent to 55 pixels or less on the camera sensor [49]. The e-SFR is orientation dependant due to the anisotropic nature of digital systems (cf. § 2.2.1). Hence, many ISO12233 test targets use edges in vertical and horizontal orientations to evaluate the e-SFR across multiple directions [35]. Horizontal and vertical orientations are typical for the e-SFR, but sagittal and tangential orientations are also commonly used (cf. § 2.6.2). The e-SFR orientations employed to quantify camera performance depend on the application and intended use.

2.3.1 System Linearisation

The tone response of digital camera systems is a function between the output pixel values and input scene luminance [50, pp. 383–386]. The RAW output tone from a system is a linear function; however, in either the camera firmware or RAW conversion software, a nonlinear response is applied, known as gamma correction. The purpose of the gamma correction is to mimic the Human Visual System (HVS) response and to account for the typical display non-linearity. When the image is displayed, the tonal values are then perceived correctly. This transfer function is known as the Opto-Electric Conversion Function (OECF).

Modern computer displays, such as Liquid Crystal Displays (LCDs), have a linear response [50, p. 381]. However, all current standards have been based on Cathode Ray Tube (CRT) displays, with an exponent power function with gamma between 1.8 and 2.4 [51, pp. 183–187]. Thus, LCDs are given a transfer function that mimics a CRT by remapping the signal via a voltage modification or programmed look-up-table (LUT) [50, p. 381].

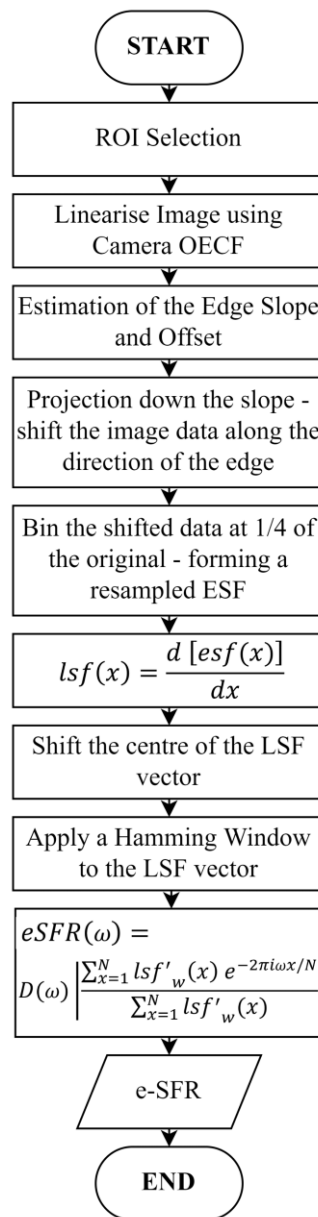


Figure 2.7 The ISO12233 slanted-edge method flowchart, adapted from the ISO12233:2017 [5].

As the MTF requires a linear system (cf. § 2.1.1), a linearisation process is necessary prior to the measurement. Linearisation is achieved by simply applying the inverse of the Gamma power function. However, camera tone reproduction curves are usually more complex than a simple power function. Therefore, in the ISO12233, the test target image is linearised by first measuring the camera's OECF through implementing ISO14524:2009 [2]. This, in turn, is used in the form of a LUT for applying the inverse to linearise the image.

The ISO14524 method cannot be implemented without knowledge of the input tone values. Therefore, implementing it with natural scenes is impractical. Linearisation is essential for

the e-SFR; without doing so, bias is introduced [52]. That said, if an imprecise linearisation is implemented, the e-SFR bias is not often such a severe problem [53]. In this work, linearisation can effectively be applied using the transfer function of the assigned colour space of the file when implementing the slanted edge method to pictorial natural scenes.

Many image formats incorporate a standardised tone reproduction embedded in the file. A typical colour space is the Adobe RGB, which has a gamma value of 2.2 [54] and only requires a simple correction process:

$$C_{RGB} = C'_{RGB}{}^\gamma \quad (2.24)$$

where $\gamma = 2.19921875$, and C is each gamma-corrected colour channel, RGB, and C' are the uncorrected RGB channels.

The sRGB is another common colour space. This colour space protocol has a more complex linearisation transfer process [55]:

IF $C'_{sRGB} < -0.04045$

$$C_{sRGB} = - \left[\frac{C'_{sRGB} + 0.055}{1.055} \right]^\gamma$$

IF $0.04045 \leq C'_{sRGB} \leq 0.04045$

$$C_{sRGB} = \frac{C'_{sRGB}}{12.92} \quad (2.25)$$

IF $C'_{sRGB} > 0.04045$

$$C_{sRGB} = \left[\frac{C'_{sRGB} + 0.055}{1.055} \right]^\gamma$$

where $\gamma = 2.4$ and the other terms are as previously defined.

2.3.2 The Slanted-Edge Algorithm

Following linearisation, the step-edge of interest from which the e-SFR is measured is cropped within a window, i.e. the region of interest (ROI), φ .

Each row, r , across the edge spread in the ROI provides an estimate of the ESF, and its derivative forms the system's discrete LSF (cf. § 2.2.1). The centroid, C , of each of the LSFs of all the rows are found by calculating the derivative per row [5]:

$$C(r) = \frac{\sum_{p=1}^{P-1} p[\varphi(p+1, r) - \varphi(p, r)]}{\sum_{p=1}^{P-1} [\varphi(p+1, r) - \varphi(p, r)]} - 0.5 \quad (2.26)$$

where P is the number of pixels and p is the pixel number per row.

The centroid from all rows is then used to find the shift of each row, $S(r)$, from a reference origin, accomplished by first calculating the slope, m [5]:

$$m = \left[\frac{\Delta r}{\Delta C(r)} \right] \quad (2.27)$$

$$S(r) = \frac{R/2 - r}{m} \quad (2.28)$$

where Δr is equal to 1 row, and R is the number of rows.

The computed centroids are used to form a projection down the slope of the input edge, shifting the data into a single ESF. This stage is represented in Figure 2.8, adapted from Reichenbach et al. [22]. Next, the shifted data is binned to form an ESF of 4-times higher resolution than that of the original image. The slope projection and subsequent binning are expressed as [5]:

$$esf'(x) = \frac{\sum_{r=1}^R \sum_{p=1}^P \varphi(p, r) \cdot \alpha(p, r, x)}{\sum_{r=1}^R \sum_{p=1}^P \alpha(p, r, x)} \quad (2.29)$$

$$\alpha(p, r, x) = \begin{cases} 1, & -0.125 \leq [p - S(r) - x] < 0.125 \\ 0, & \text{otherwise} \end{cases}$$

where the α function is a counter and switch to include or exclude a value in any bin.

This resampled ESF is then differentiated (Equation 1.2) and windowed to obtain the LSF [5]:

$$lsf'_w(x) = W(x) \frac{d[esf(x)]}{dx} \quad (2.30)$$

where x is the spatial distance in pixels, and the window is expressed as:

$$W(x) = 0.54 + 0.46 \cos[2\pi(x - 2X)/4X] \quad (2.31)$$

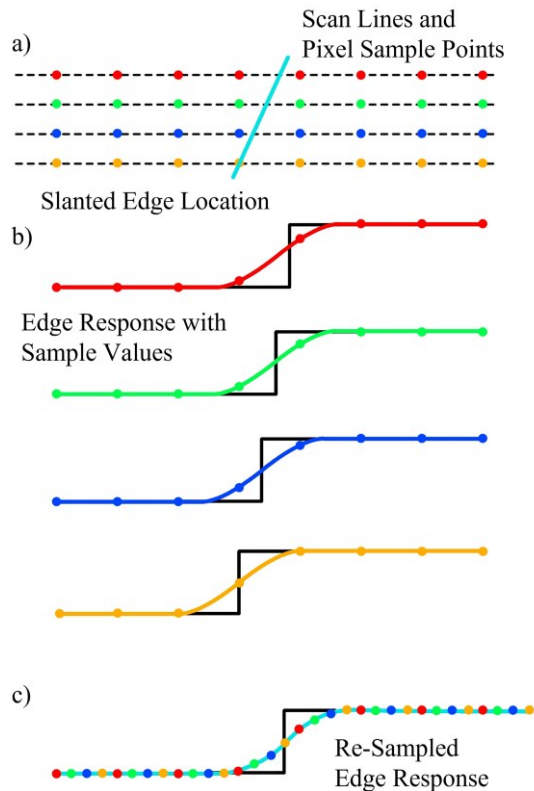


Figure 2.8 Registering scan lines with a slanted edge, allowing the ESF to be resampled. Adapted from [22].

The modulus of the DFT of the LSF is calculated to yield the e-SFR [56, 23] (Equation 1.3). In this calculation, the modulus vector is normalised by the e-SFR at the zero-frequency [5, 57]. The e-SFR is typically corrected for the discrete derivative response, $D(\omega)$ [5]. This is expressed as:

$$eSFR(\omega) = D(\omega) \left| \frac{\sum_{x=1}^N lsf'_w(x) e^{-2\pi i \omega x / N}}{\sum_{x=1}^N lsf'_w(x)} \right| \quad (2.32)$$

2.3.3 Benefits of the Slanted-Edge Method

The slanted edge method provides advantages over other performance measurements. First, resampling of the ESF allows the e-SFR to extend beyond the Nyquist frequency without the influence of aliasing [58]. Secondly, the resampled ESF from many rows from a ROI removes the non-stationary system effects of the system (cf. § 2.2.3), obtaining the in-phase pixel response [59]. Also, if required, the slanted-edge method can be adapted to obtain the out-of-phase e-SFR, as proposed by Jenkin et al. [60].

The target simplicity is beneficial for the practical implementation of the method [56]. Any high quality printed step-edge can be used effectively, making a step-edge test chart cost-effective and straightforward to create and print. In addition, unlike the original edge method, the slanted-edge method does not need to be aligned with the sampling array, providing a practical application for digital systems. In fact, it must not be aligned; the misalignment is required for the resampling to occur.

2.4 Sources of e-SFR Variation

The slanted-edge method has been shown to be a repeatable and robust measure of performance. However, sources of variation in the measurement should be controlled to evaluate the e-SFR correctly. Such sources include a number of input parameter values, including edge angle, contrast, and ROI window size (cf. § 2.4.1), the effects of the system ISP (cf. § 2.4.2), and the performance variation introduced by the optical system (cf. § 2.4.3).

2.4.1 Edge and ROI Parameter Ranges

Edge contrast has little effect on the e-SFR, provided that the system to be measured is nearly noiseless with a linear ISP. Low contrast edges become more corrupted with noise than high contrast edges [61] (cf. § 2.5.1), whilst high contrast edges are more prone to non-linear sharpening [28, 62] (cf. § 2.4.2). Consequently, the standard recommends a low to medium Michelson contrast of 0.55 to 0.65 to minimise the effect of these conflicting negative attributes [5].

Further, the ISO12233 recommends a 5° edge angle with respect to the orientation of the ROI [5]. Previous studies showed that increasing edge angle increases the measurement error [61, 63, 64], but generally, a 3° to 30° angle range from the vertical has been found acceptable [63]. There are problematic angles that need to be considered when using the slanted-edge algorithm. These critical angles are caused by ROIs that have a specific number of rows combined with edge angles that create zero-count bins in the oversampling process [65–67]; in other words, missing values in the resampled ESF. Van Den Bergh provides simulated data that illustrates such critical angles when using an 8-times bin and a ROI height of 30 pixels [67], as seen in Figure 2.9. This data shows that several angles do not meet the required oversampling factor. Lowering the number of bins reduces the number of critical

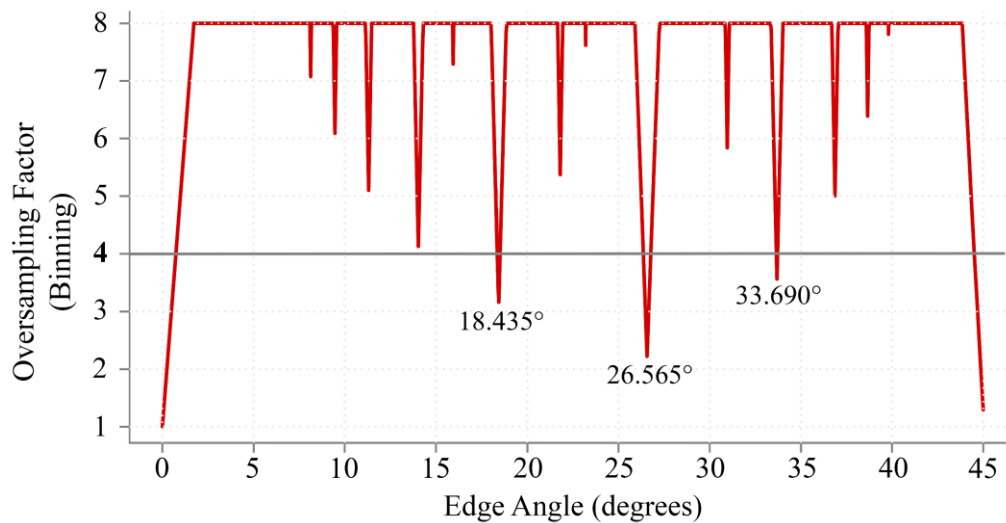


Figure 2.9 The mean oversampling factor, capped at 8-times, for simulated step-edges ranging from 0 to 45 degrees. The ROI window height used was 30 pixels. Adopted from [67].

angles; the 4-times bin is marked in this figure, with 18.435° , 26.565° and 33.690° not meeting this oversampling factor. Using taller ROI than the 30 pixels further reduces the number of critical angles. For the ISO12233, if the ROI window size is appropriately selected, see below (Table 2.1), critical angles are not usually problematic. That said, when using natural scenes, both edge angle and height are both dependent on the pictorial image. For instance, busy scenes may require short ROI heights to isolate a step-edge.

The ROI window size also introduces variation in e-SFR [63, 64]. The vertical ROI (height) determines the number of ESFs used in the edge resampling process. More rows lead to more accurate results (cf. § 2.5.1). Generally, it is recommended to be greater than 80 pixels for the highest accuracy, but it can be decreased to 40 pixels if noise is low [68].

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. With high image noise, a wide ROI will include more of the noisy signal on either side of the step-edge, introducing high-frequency bias to the e-SFR (cf. § 2.5.1). Hence, a narrow ROI is recommended [63, 69], as long as it does not inhibit the ESF. In addition to a small width, a tall ROI height keeps the ROI vertical-to-horizontal aspect ratio high [63], diminishing the effects of noise on the resulting e-SFR.

Table 2.1 lists the standard parameter ranges recommended by the ISO122233.

Parameter	ISO12233 e-SFR
Edge Angle	<45° (5° Recommended)
Edge Contrast	0.55 – 0.65 Michelson Contrast
ROI Size	>64 x 80-500 pixels

Table 2.1 ROI and edge parameter ranges recommended for the ISO12233 slanted-edge method.

These standard parameter ranges are restrictive and limit the number of edges that can be extracted from a pictorial natural scene. Subsequently, the parameter ranges employed in a natural scene adapted slanted-edge SFR method require further investigation in this thesis (cf. § 4.2).

2.4.2 Non-Linear Image Signal Processing

The SFR/MTF is based on linear system theory (cf. § 2.1.1); therefore, problems occur when measuring the performance of the non-linear systems. The non-linear effect on the SFR/MTF is not a new issue; it is present in all implementations of linear system performance measurements. The chemical transfer function [10, p. 239, 70] is an example of system non-linearity caused by the chemical process of photographic film image formation. With the introduction of digital systems and the current miniaturisation development, non-linear processing has become conventional in most modern systems, significantly improving image quality but increasing the complexity of deriving an SFR/MTF. This issue still requires further research as many standards proposed for measuring non-linear systems continue to use methods based on linear system theory, such as the IEEE P1858 CPIQ standard, developed for smartphone systems [71], as well as the ISO12233 [5].

The e-SFR is particularly susceptible to gradient-based sharpening, where the sharpening strength is proportional to the edge gradient [20, 72]. Consequently, when a non-linear ISP is incorporated into the system pipeline, the measured e-SFRs become signal-dependent. Thus, to minimise the effects of non-linear sharpening on the e-SFR, ISO12233 recommends utilising a medium-contrast edge (cf. § 2.4.1).

Jenkin [20, pp. 162–165, 72] proposed a method to estimate an SFR based on the idea of a *contrastless* edge. The theory behind this approach is to obtain the SFRs of a series of edges

at various contrasts, then to use interpolation to estimate the SFR at zero contrast, giving an approximate SFR of the system without processing. Such a method assumes that there is little noise corrupting the high frequencies of the low contrast edges (cf. § 2.5.1).

Linear denoising reduces the performance of the system as it adds blur globally. In contrast, when measured from a test chart, non-linear denoising does not significantly affect the e-SFR. This is because it adds blur locally, mainly in the uniform noisy areas, maintaining the sharpness of strong edges. Isolated test chart edges with uniform tones on either side make it straightforward to preserve the edge. With a natural scene derived SFR, the non-linear denoising is more problematic. This is because complex pictorial natural scenes, captured with a non-linear ISP, contain edges that are processed depending on the gradient, surrounding image noise and scene textures, resulting in various local scene dependant PSFs.

2.4.3 Optical Performance Variation

Camera performance varies across the field of view due to restrictions in the optical system. The highest performance is obtained in the central region of the optics, whilst the frame corners have reduced sharpness and resolution. This performance drops off due to increased aberrations, including coma, astigmatism, field curvature, and distortion, which worsen with increasing field angle and distance off-axis [73, pp. 82–83]. The extent of this performance variation depends on the characteristics of the optical design. The performance variation across the imaging circle may have performance peaks and troughs, i.e. it is not necessarily a linear degradation from the centre to corners of the field of view. As a result, any performance assessment should be related to the edge location in the frame.

Figure 2.10 illustrates this variation for a wide-angle lens; e-SFRs were measured from an ISO12233 test chart, cropping many 64x128 pixel ROIs across the entire field of view. The resulting e-SFRs are colour coded to depict the edge location as a radial distance from the centre to the corners of the frame.

The ISO12233 does not describe how this variation should be depicted or how to combine measurements across the frame, as it is application dependent. The e-SFRs, or a metric based or derived from the e-SFRs (cf. § 2.6.1), can be averaged using radial zone weights, for instance, the weighted averages employed in the Imatest software [74] or the ISO20462 IQM [75]. Alternatively, the e-SFRs can be plotted against radial distance, such as sagittal and tangential graphs (cf. § 2.6.2).

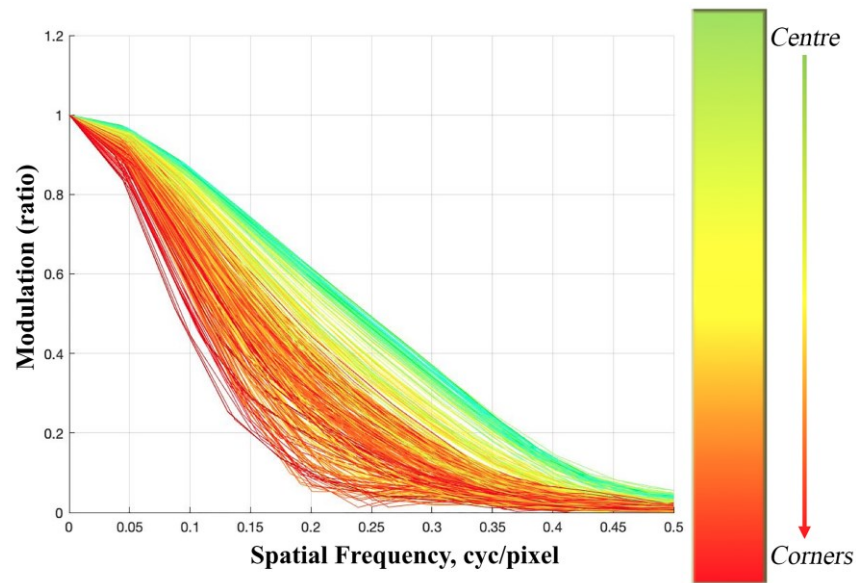


Figure 2.10 Illustration of the e-SFR variation introduced by edge location within the field of view, captured using the *Nikon D800* with a 24mm lens at $f/4$ [76].

2.5 Sources of e-SFR Error

Several sources of error in the slanted-edge method need to be considered when implementing the method. Image noise (cf. § 2.5.1), ROI non-uniformity (cf. § 2.5.2), and clipping (cf. § 2.5.3) are all sources of error that are controlled in the ISO12233 but need to be evaluated when adapting the method for natural scene inputs.

2.5.1 Image Noise Corruption

Image noise corrupts the e-SFR and is the most significant source of error, introducing positive bias to the high-frequencies [39, 56, 63]. The source of image noise is commonly from the system (cf. § 2.10); thus, in practice, the image is photographed using conditions to minimise system noise, including a well-illuminated test chart with low ISO gain. High-frequency scene textures surrounding the edge are a second noise source when utilising natural scenes for SFR determination.

Figure 2.11 demonstrates high-frequency image noise corruption. Eight ROIs were simulated with identical edges with decreasing Signal-to-Noise Ratio (SNR) values [4] (cf. § 2.10). Image shot and read noise were simulated for the purposes detailed in Section 4.2.3.

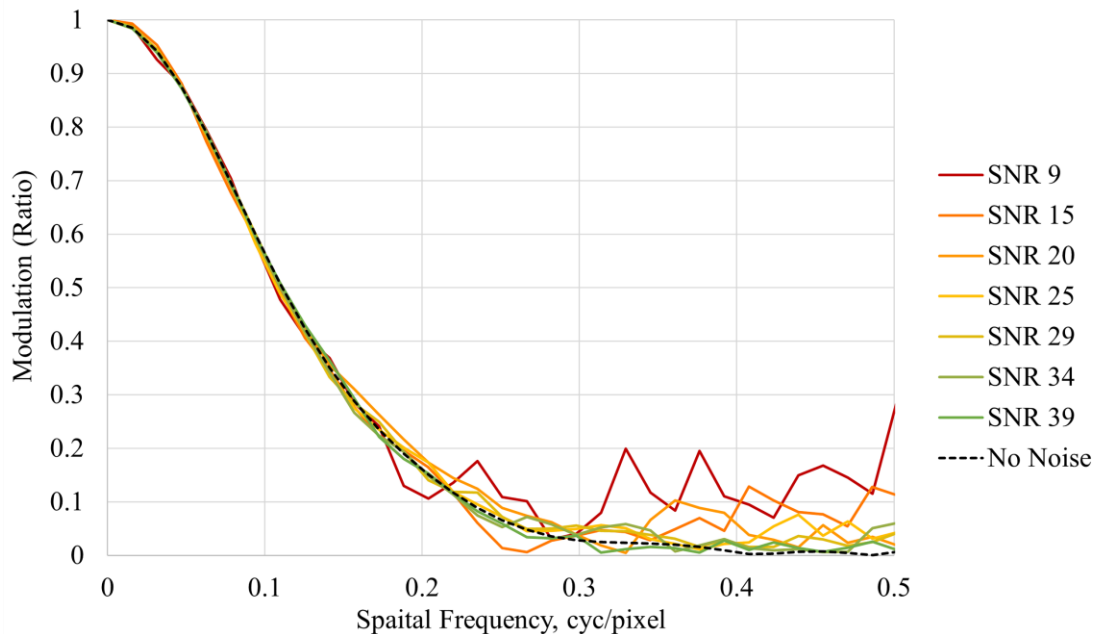


Figure 2.11 The output e-SFRs for eight simulated ROIs with decreasing SNR values.

Averaging multiple e-SFRs together helps remove random noise [39]; this should be done in the spatial domain, i.e. averaging ESFs or LSFs. If SFRs are averaged in the frequency domain, the bias accumulates rather than being reduced. Also, this bias can be diminished by isolating tall ROIs in the slanted-edge method. Incorporating more ROI rows in the resampling of the ESF (cf. § 2.3.2) allow random noise to be averaged out (cf. § 2.4.1).

Yeadon et al. [39] proposed a confidence limit method that utilises the Phase Transfer Function (PTF) to establish precision for a single measurement. The PTF is zero or linearly related to the spatial frequency for a noiseless edge input. Any non-linear PTF behaviour can be interpreted as noise corruption in the edge trace.

The Yeadon et al. confidence limit methodology is as follows [39]:

1. If the area under the PTF is a non-zero value, apply a correction to ensure a zero area. First, use values up to $\frac{3}{4}$ of the cut-off frequency, k_c , to calculate the correction.
2. Use the modulus component, M , and modified PTF, ϕ , to calculate the real, R , and imaginary, I , components of the OTF:

$$\begin{aligned}
 R(\omega) &= M(\omega) \cos \phi(\omega) \\
 I(\omega) &= M(\omega) \sin \phi(\omega)
 \end{aligned}
 \tag{2.33}$$

3. The imaginary component represents a cone-shaped error projection, Figure 2.12. The semi-angle, α , of this error projection is then calculated for each spatial frequency up to the termination frequency, ω_t :

$$\alpha(\omega) = \frac{\tan^{-1} I(\omega)}{\omega} \quad (2.34)$$

4. The root-mean-square value (RMS), α_{rms} , of the semi-angle function is then calculated:

$$\alpha_{rms} = \sqrt{\frac{1}{\omega_t} \int_0^{\omega_t} \alpha^2(\omega) d\omega} \quad (2.35)$$

5. Steps 1 to 4 are repeated between values $\frac{3}{4}$ to $\frac{1}{4}$ of the cut-off frequency.
6. Select the lowest RMS value to provide the confidence limits:

$$eSFR(\omega) = R(\omega) \pm C(\omega) \quad (2.36)$$

$$\text{where } C(\omega) = \pm \omega \tan \alpha_{rms}$$

The error increases with spatial frequency. These confidence limits represent statistical limits with a 70% probability of including the actual value. Figure 2.12 a) illustrates the possible spread from the OTF, and b) the MTF.

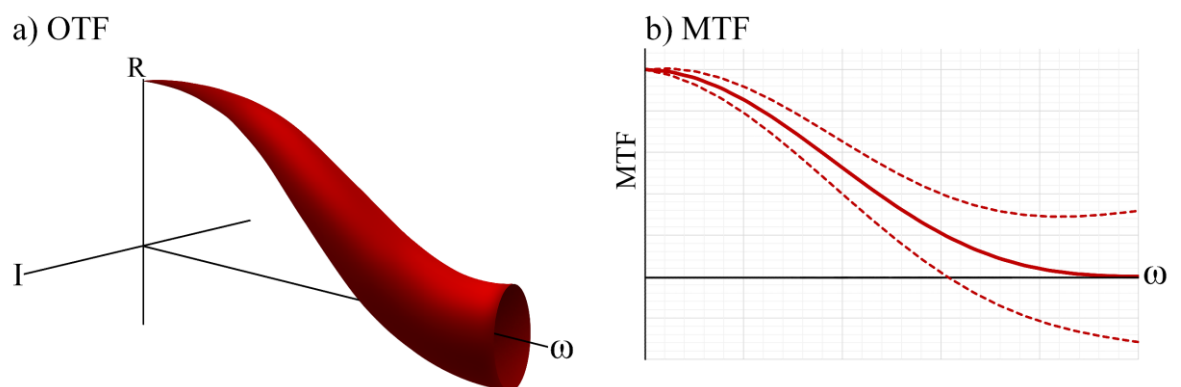


Figure 2.12 The confidence limits depicting the possible spread due to noise corruption from a) the OTF and b) the MTF. Adopted from [39].

2.5.2 Tonal Non-Uniformity

The ISO12233 states that the test chart must be photographed under controlled lighting conditions, with a maximum illumination delta of 2% (cf. § 2.3) [5]. Non-uniform lighting causes low frequency gradients through extracted ROIs. In turn, this distorts the ESF profile and introduces e-SFR bias.

In natural scenes, lighting and tone are rarely uniform; as a result, many natural scene ROIs contain illumination non-uniformity. Unless corrected, this causes a significant source of error in natural scene derived SFRs.

The non-uniformity correction employed in the Imatest software [77] provides a simple solution. This method fits a first-order function to the light side of the ROI, as the lighter pixel values inherently contain a higher SNR. As this area is meant to be of uniform tone, if the fitted function has a gradient, the inverse of this function is applied to provide a corrected ROI. For test chart images, this method works well and is due to be part of the upcoming revision to the ISO12233:2022 [78], due later this year (2022).

This non-uniformity bias is depicted in Figure 2.13, where a 128x128 pixel ramp-edge ROIs, i.e. gradient step-edge, was simulated using a hyperbolic tangent (*tanh*) function (cf. § 4.2.1), Equation 4.1. Five different angled low-frequency gradients were added to the simulated ROI, resulting in five tonal non-uniformities. These simulations were used to demonstrate the effect of non-uniformity on the e-SFR.

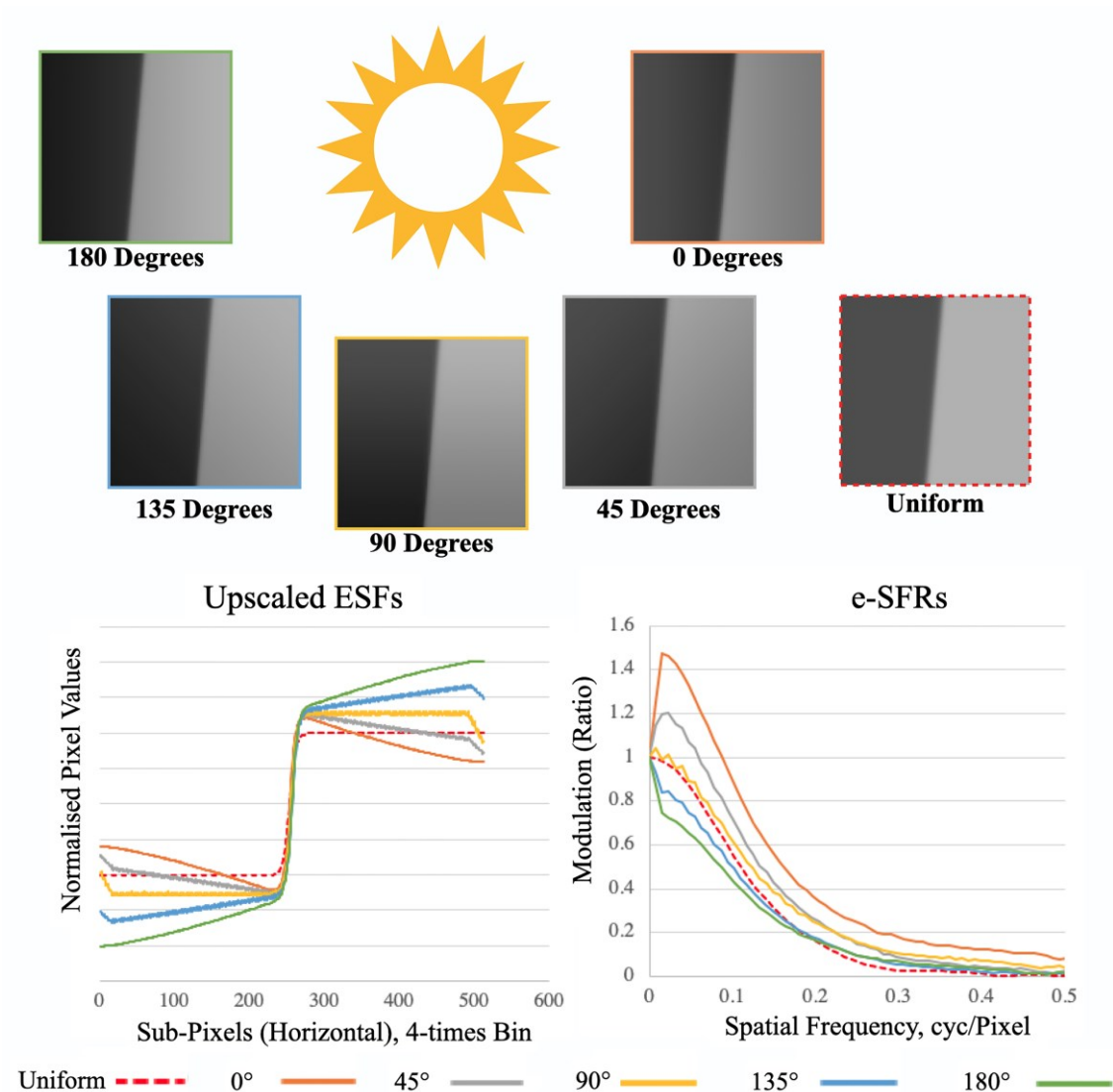


Figure 2.13 The bias introduced to the e-SFR with low frequency non-uniformity. The ESF and SFR demonstrate how these non-uniformities affect the result compared to a uniform ROI. Adapted from Imatest [77].

2.5.3 Signal Clipping (Saturation)

When a signal exceeds the dynamic range of the system, it is clipped. The system response at these saturated values is considered non-linear [63, 69]. When the step-edge tone just reaches the maximum or minimum value, the e-SFR is unaffected; this is referred to as soft-clipping. Severe clipping creates a hard shoulder in the ESF, which in turn overestimates the system's e-SFR. The effect on the e-SFR with soft and hard clipping is demonstrated in Figure 2.14. These ESFs were created using noiseless simulated ROIs.

In practice, soft-clipping causes error in the e-SFR, as pixel values from noise fluctuations and sharpening effects are saturated [69]. Hence, when implementing the slanted-edge method, the exposure should be set so that the target values fall around the centre of the system's dynamic range. A low contrast target also helps to avoid clipping [79].

Natural scenes have a broad range of step-edge contrasts, some of which would be clipped. Such edges should be detected and removed from the evaluation (cf. § 3.5).

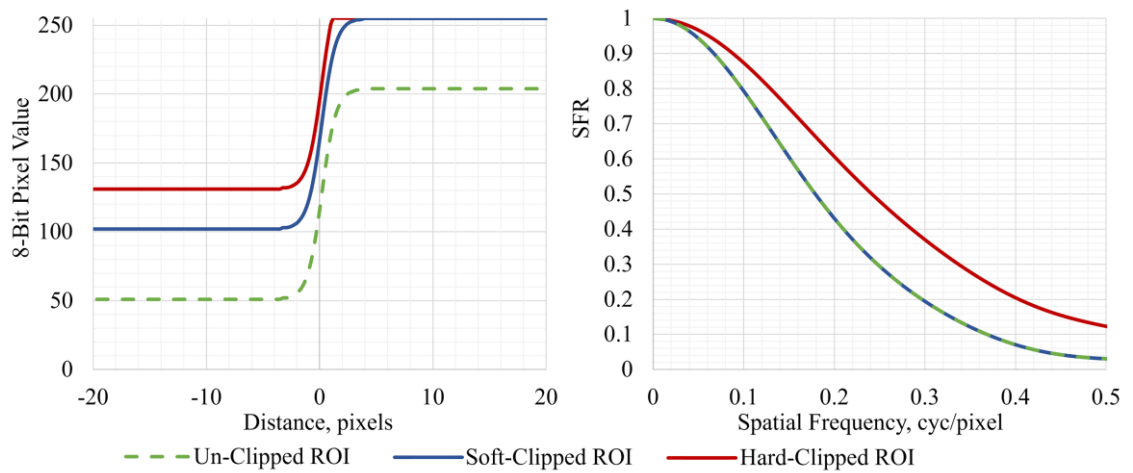


Figure 2.14 Simulated noiseless ROIs at a 0.60 Michelson Contrast with mid-range un-clipped pixel values, soft-clipped pixel values, and hard-clipped pixel values. The output ESFs and the resulting e-SFR are plotted, illustrating the clipping error.

2.6 Assessing e-SFRs

2.6.1 Comparison Metrics

Comparing and assessing system performance is achieved through taking specific metrics from the SFR/MTF. Taking a single value from a function does not provide the whole story with regards to system performance but helps compare systems or the method's accuracy.

Due to noise corrupting high frequencies, the low-frequencies tend to have the highest accuracy (cf. § 2.5.1); therefore, comparison metrics are often taken in the mid to high frequencies, where the error is significant. The most common being the MTF50, MTF20 and MTF10, i.e. the frequencies that correspond to 50%, 20% and 10% of the modulation, respectively. In addition, these metrics also convey important system performance characteristics. The MTF50 is a measurement of sharpness, whilst the MTF10 provides the

system limiting resolution [6, p. 449], i.e. the spatial frequency where two points are just distinguishable, as defined by the Rayleigh criterion. However, it should be noted that these metrics do not account for image noise [58]. When using MTF20/10, the corruption effects of image noise will be most significant. Consequently, the boosted high frequency may misrepresent the resolution capabilities of the system. In addition, for extreme noise, the boosted high frequencies may result in no MTF10 metric below the Nyquist frequency.

The MTF50/20/10 metrics also are susceptible to sharpening. As a result, these metrics incorrectly represent a system that has been subject to sharpening, overestimating the actual performance. Koren [49, 50] proposed the MTF50P and a normalised area metric as a solution. The MTF50P is similar to the MTF50, but is 50% of the peak modulation. The normalised area is the area under the MTF curve (up to the Nyquist frequency) normalised to a peak value of one. Both the MTF50P and the area provide a consistent metric with an increase of sharpening [80]. Nevertheless, the MTF50P is preferred over the normalised area, as it is a more familiar metric in practice and better represents the MTF50.

An alternative approach is to take the modulation at specific frequencies. Comparing SFRs/MTFs at set frequencies illustrate the systems' rendering ability at the same level of detail. For example, this could be at 10, 20, 30, 40 cycles/mm or for a set pixel size at 25%, 50% and 75% of the Nyquist frequency. Depending on the application, this may be a more helpful metric.

Combining such metrics to calculate the performance variation across the field of view can provide an insightful comparison between systems. An example is to measure the asymmetry of a lens [81]. Due to misaligned optical elements, or the lens and sensor not being perfectly parallel, the performance may not be symmetrical across the field of view. The system asymmetry is measured using metrics, such as the MTF50, in four frame locations. Measurements are taken from the right, left, top and bottom of the frame. Imatest obtains the MTF50 by fitting a second-order function of the measured data and using this function to obtain a measurement for each frame location [81]. The x-axis and y-axis asymmetry is calculated by [81]:

$$\begin{aligned} \text{MTF asymmetry}(x) &= \frac{MTF50R - MTF50L}{MTF50R + MTF50L} \\ \text{MTF asymmetry}(y) &= \frac{MTF50T - MTF50B}{MTF50T + MTF50B} \end{aligned} \tag{2.37}$$

The sampling efficiency is another helpful metric [5, 58]. It is a measurement of the likely signal detail derived from the system, expressed as a percentage. This measurement is calculated from the limiting resolution, R_L , and the sampling resolution, R_S . As mentioned above, for digital systems the limiting resolution is given by the MTF10. The sampling resolution is half the sample rate, i.e. the Nyquist frequency. The 1D sampling efficiency, E_{1D} , is expressed as [58]:

$$E_{1D} = \frac{R_L}{R_S} \cdot 100\% \quad (2.38)$$

Obtaining the sampling efficiency for the vertical, horizontal and diagonal e-SFRs, the 2D sampling efficiency, E_{2D} , is calculated by [5, 58]:

$$E = \frac{E_d(E_V + E_h)}{2} \quad (2.39)$$

2.6.2 SFR Orientation

As discussed (cf. § 2.3), digital systems are anisotropic, meaning performance is direction dependant. It is common to measure the e-SFR using horizontal and vertical edges. The measured e-SFR is perpendicular to the edge; in other words, a vertical edge with respect to the sampling array provides the horizontal e-SFR.

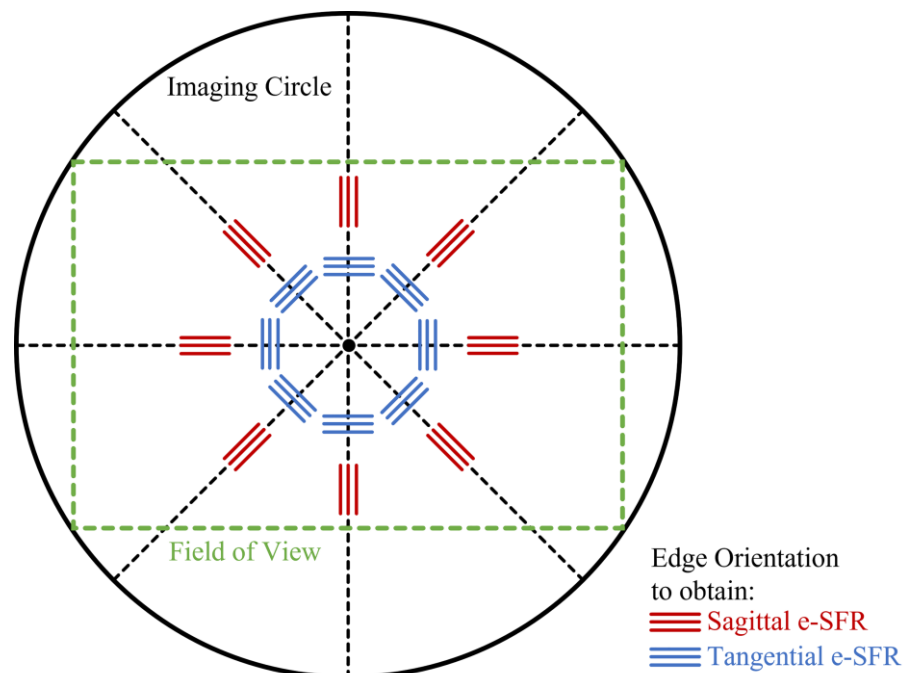


Figure 2.15 Illustration of the edge orientations for sagittal and tangential e-SFR analysis in relation to the optical imaging circle and field of view.

The performance of the optical system also changes across its projected imaging circle (cf. § 2.4.3). The slanted-edge method can be used to provide horizontal and vertical e-SFRs across the field of view. That said, more appropriate edge orientations for assessing performance against radial distance, particularly for optical system analysis, are sagittal and tangential (also referred to as radial and meridional) [15, 82]. Figure 2.15 shows the edge orientations for sagittal and tangential e-SFR measurement; they are related to the radial orientations of the imaging circle and the anisotropic PSF as it moves away from the optical centre.

The sagittal and tangential e-SFRs/MTFs are commonly plotted using the modulation value at a stated metric, either MTF50/20/10, the 25%/50%/75% of the Nyquist frequency, or similar, which is plotted against radial distance. In many cases, several of these metrics are plotted to summarise the system performance. This process is depicted in Figure 2.16, adapted from [15]. The sagittal and tangential e-SFR metrics are plotted against the diagonal field of view position from corner to corner, crossing the centre. Alternatively, the position data points can be averaged, providing an azimuthal e-SFR per radial distance. The azimuthal average is used in this thesis, as it provides more data from the entire frame for establishing a natural scene derived SFR.

Using the ability to measure the sagittal and tangential e-SFRs in the natural scene derived performance methodology allows for many optical analysis investigations, including assessment of the astigmatism of the lens. Astigmatism is when the sagittal and tangential focus regions do not align, the distance between them is known as the astigmatic difference

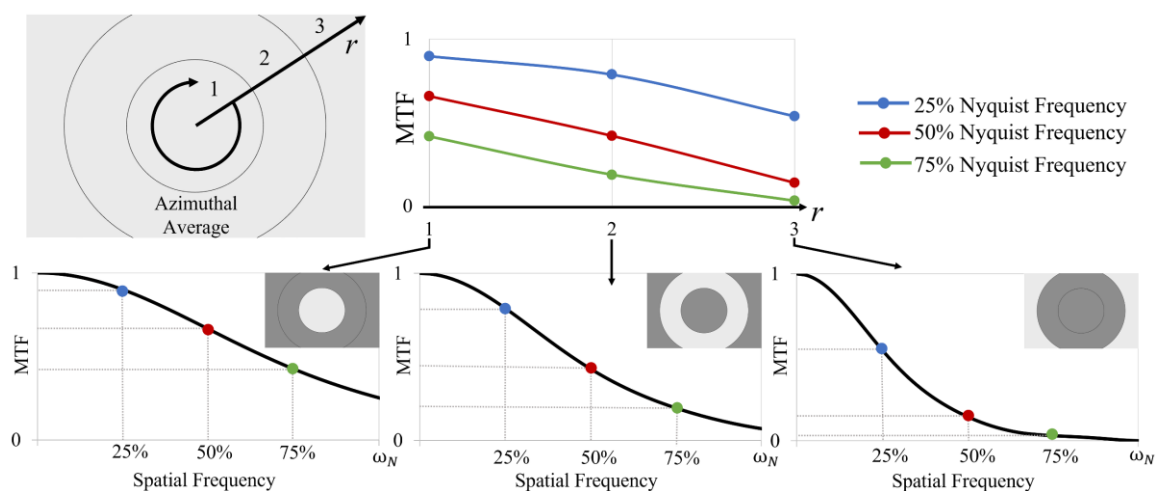


Figure 2.16 The process of plotting the sagittal and tangential e-SFRs/MTFs for 25%/50%/75% of the Nyquist frequency and how it relates to the traditional MTF bell curve. Adapted from [15].

[73, pp. 89–90]. It occurs when an off-axis point is not sharp, seen commonly in the corners of the field of view, and is detected when the sagittal and tangential e-SFRs do not match across the frame, i.e. one becomes significantly worse than the other.

2.7 Improvements to ISO 12233

2.7.1 General e-SFR Developments

The ISO12233 is consistently being improved; it has had three published iterations from its first release in 2000. Many studies have proposed improvements to the standardised slanted-edge method since its last iteration in 2017, aiming to increase the accuracy and precision of the measured e-SFR. Some of these will be implemented in the upcoming fourth iteration of the standard, the ISO12233:2022 [78].

The ISO12233:2017 standard uses a linear fitting function to determine the edge slope. Straight input edges may become curved in the image due to optical geometrical distortions, which results in e-SFR bias. This bias can be avoided using a higher-order polynomial fitting function [83–86]. Under simulated testing, the result from a third-order polynomial fitting was almost identical to linear-fitting when the edge input is straight [86]. When a distorted edge was used, the polynomial fit returned better accuracy. Since 2018, a higher-order polynomial fitting function has been considered for the ISO12233 standard [86]. It has now been announced that a 5th order polynomial fitting function will be part of the ISO12233:2022 [78]. Such fitting functions have already been implemented in industry software, including Imatest [74] and *sfrmat4* [87]. Besides the geometrical distortion correction, higher-order fitting is valuable for natural scene edges, which are rarely straight.

Baer has taken the slanted-edge method a step further to derive an e-SFR from a circular edge [85], achieved through a similar approach as the higher-order polynomial fitting function. Instead of a function, radial scan lines were fitted to a circle, described by a Fourier series. This allows the e-SFR to be measured from the circular edge with results equal to the standard.

Williams and Burns [83] have proposed a filter that helps decrease the effects of image noise on the e-SFR. Referred to as the filtered tails procedure, it applied a Gaussian blur on either

side of the edge while preserving the ESF. The blur removes the variance in the uniform areas of the ROI, resulting in the measurement of e-SFR without significant noise bias.

Another proposed modification to the slanted-edge method is the multi-directional MTF estimation by Masaoka et al. [84]. The aim of this method was also to improve the e-SFR estimation from images with high noise. It was achieved by modifying how the edge position is found. In this method, a two-dimensional function was fitted to the ROI instead of the standard method of taking the derivative of each pixel row. This modified method was tested against the standard slanted-edge method by measuring the e-SFR 1000 times for a 5° angled edge with SNRs ranging between 20dB and 60dB. As a result, this proposed method was more precise and robust against noise than the ISO12233 e-SFR.

Birchfield proposed the Reverse-Projection method [26] to reduce the ISO 12233 slanted-edge standard variation, caused specifically by edge angle and the size of the ROI. To accomplish this, Birchfield proposed significant changes to the ISO12233. These changes start with calculating the centroid points within the pixel rows, the edge fitting parameters and the resampling processes. The standard method computes the centroid through derivatives along the rows, then calculates the slope of the edge through a shift. The ESF is created through a forward projection down the edge slope (cf. § 2.3.2). The Reverse-Projection method proposal changes these three steps. A Ridler-Calvard binary image segmentation algorithm was applied to the ROI to segment the edge into the light and dark regions. The result yields the coordinates of the edge running through the ROI. Instead of finding the edge slope, the line that best fits the coordinates was calculated using the total least squares technique [88]. Each pixel of the resampled ESF was obtained by an average

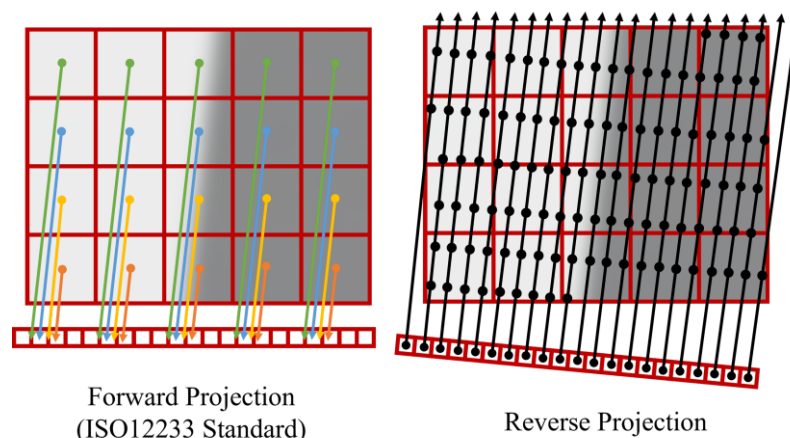


Figure 2.17 A visual representation of the reverse projection method compared to the standard forward projection method. Adopted from [64].

linearised pixel value, calculated by reverse projecting into the interpolated two-dimensional array. Visually the forward-projection (i.e. the standard) and the reverse-projection are shown in Figure 2.17 [64].

The Reverse-Projection method is unaffected by the size of the ROI or the edge angle. In addition, the standard method adds high-frequency noise to the measure when the ESF is super-sampled, resulting in an e-SFR that can be overestimated. Whilst the reverse-projection yields stable results with any amount of upsampling. This method would be advantageous in natural scene derived SFRs, as natural scenes contain edges with various parameter values. The results presented were shown to be underestimated compared to other standard implementations of the slanted edge method [64].

Imatest has proposed an alternative approach to correct edge angle variation [61]. The LSF was scaled to correct for the edge rotation. The width of the LSF perpendicular to the edge, d , is mathematically expressed as [61]:

$$d = l \cos \theta \quad (2.40)$$

where l is the width of the LSF across the pixel array and θ is the edge angle.

The LSF was then scaled to correct for this rotation [61]:

$$lsf_{corr}(x) = lsf(x \cos \theta) \quad (2.41)$$

As a result, the angle variation introduced in the e-SFR was significantly reduced. This LSF scaling correction will be added to the upcoming revision of the ISO12233 [78].

2.7.2 Real-Time MTF

Masaoka proposed real-time (60 frames per second) MTF measurements for broadcasting camera systems [66]. This method is based on the ISO12233 e-SFR methodology, achieved using a starburst test chart containing multi-directional step-edges, shown in Figure 2.18. It employs a test chart; thus, it is carried out under controlled conditions.

To summarise this method, the first input video frames are averaged to calculate the expected noise level, which is used to reduce the bias introduced to the measurement. The resampling works with a similar concept to the Reverse-Projection method [26] (cf. § 2.7.1), orienting the sub-pixel sampling perpendicular to the edge slope; an 8-times bin is employed here. The

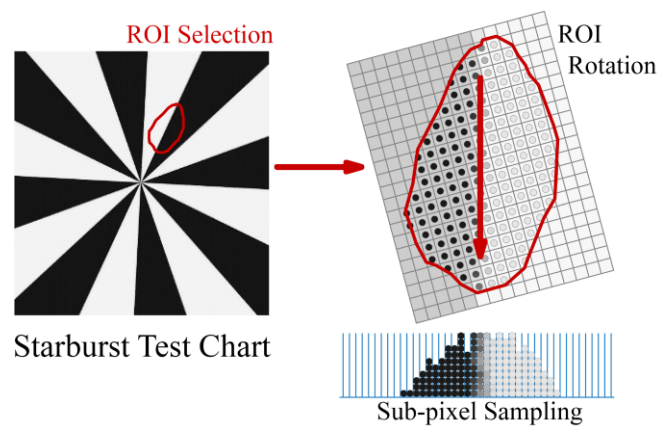


Figure 2.18 Selection process for Masaoka's real-time MTF measurement. Adapted from [66].

pixel values are obtained by averaging the values that fall into each sub-pixel. This allows an average ESF to be obtained, which is employed to measure the MTF using the established spread function calculations (cf. § 2.2.1). Following this, assuming that the edge angle stays constant, a look-up-table (LUT) is created to establish the map of the projection of the ROI edge slope into the bin locations. Using a LUT for this task significantly reduces the computation cost of resampling the ESF, allowing real-time measurements from a single ROI, sized 100x200 pixels, at 60 frames per second.

This proposed approach was shown to have several benefits. As with the ISO12233, a step-edge test chart is simple to obtain and implement. The MTF can be measured from a small ROI that does not have to be rectangular, allowing for the natural variation of an optical system to be measured. The frame averaging and proposed binning process made the data robust against image noise. As the MTF can be extracted from a live feed, it can be used to quickly assess the performance across several focal lengths of a zoom lens over multiple apertures, determining the ideal optical settings for the system. In addition, it can be used to find the objective optimal focus and calibrate the ISP. Also, the real-time measurements allow de-focused MTFs to be quickly obtained, which is useful for astigmatism analysis (cf. § 2.6.2). Unlike the traditional e-SFR, this measurement is unaffected by edge angle. Thus, multi-directional MTFs can be extracted, enabling direct measurements of the system anisotropy (i.e. misalignment between the optical and sensor planes).

2.8 Natural Scene Derived System Performance

Measuring the SFR/MTF from natural scenes is not a totally new concept. Previous studies have used methods such as the slanted-edge algorithm (cf. § 2.8.1), the texture-MTF algorithm (cf. § 2.8.2), Blind Deconvolution (cf. § 2.8.3) and convolutional neural networks (CNN) (cf. § 2.8.4) to either correct or estimate the system performance directly from pictorial natural scenes.

There is a drawback when obtaining the system performance from scene inputs, affecting all methods detailed below. If the image does not contain sufficient content to employ a specific performance measure, it cannot be implemented. This same caveat applies to the method proposed by this thesis.

2.8.1 Edge Isolation

Previous studies have used natural scene step edges to assess aerial/remote imagery [89–93]. A satellite's SFR is often characterised using traditional laboratory methods before launch. The launch process, going from air to a vacuum, with thermal changes to the sensor and imaging through the atmosphere, changes the output system performance compared to the laboratory-based performance [90]. Therefore, performance should be monitored when the satellite is in orbit for calibrating the images to the expected SFR. Many applications have been proposed to accomplish this task, one of which utilised step-edges to measure the e-SFR. These step-edges were extracted from captured large targets positioned in known locations [89, 93]. The disadvantages of such targets were the required size and that the measurement can only be taken when the satellite is in a specific orbit location. Alternatively, images of urban areas, buildings, roads, agricultural field boundaries would provide suitable step-edges for the e-SFR [89, 90, 92]. Another solution utilised the moon's edge to obtain a lunar MTF [85], which can provide successful calibration. The edges can be selected manually or through an automated process utilising edge detectors [89].

Another application optimises digital scan resolution for film archives [94]. The proposed algorithm used photographic negative film slides scanned at high resolutions, for example, 5000 pixels per inch (PPI). The edges in these scans were located using the Canny edge detector [95], the Hough transform was then used to select straight edges, and ROIs were isolated. The selected ROIs underwent a selection process to obtain edges with an angle of

at least 3° and a contrast level of at least 20%. Any overlapping edges were removed before using these ROIs as the input to the slanted-edge algorithm. The optimal scanning resolution was determined by applying the three-sigma rule, using the mean distribution of sampling efficiencies (calculated using Equation 2.38 and the standard deviation. The result was shown to be stable and independent of the initial scanner settings [94].

These approaches present robust methodologies for using natural scene step-edges as a qualitative assessment. The film archive optimisation resembles the project's proposed method, utilising similar processes (cf. § 3.1). However, neither of these applications were developed to acquire reliable performance measures from natural scenes compatible with the e-SFR standard measurement. Instead, they were mainly designed to assist with calibration.

2.8.2 Signal-and-Process-Dependent MTFs

The texture-MTF method [27, 29] (cf. § 2.1.4) was employed with natural scene images to provide a signal-and-process-dependent MTF (SPD-MTF) [12, 13, 96] that account for the power spectra of natural scenes and the effect of non-linear, adaptive ISP response of the system. SPD-MTFs were proposed for assessing the impact of non-linear ISPs. Such measurements can improve the performance of relevant signal-to-noise IQMs [13, 17].

The texture-MTF method traditionally uses test charts with a known input signal and noise spectra, such as dead leaves [27] or spilled coins [32] charts. 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 the input signal [12]. Subsequent studies used camera pipeline simulations to provide control over the signal and noise spectra variations, as well as ISP variants [13]. For applications outside the laboratory, the natural scene signal and noise spectra are inaccessible; thus, the texture-MTF is not appropriate to achieve the aims of this thesis.

2.8.3 Blind Deconvolution

When studying the estimation of performance from images of natural scenes, the process of Blind Deconvolution is commonly suggested, as it is a well-established procedure to estimate and correct for system blur.

The camera system MTF can be calculated directly from the PSF [18], as stated (cf. § 2.1.4 & 2.2.1), the system performance is expressed as the DFT of the input function, convoluted by the DFT of the sampling comb (the system's PSF) (see Equation 2.6). This concept is rarely used as it is impractical to obtain an impulse of light narrow enough. The concept of Blind Deconvolution was first introduced in the 1990s by Haykin [97, 98]; it refers to the process which attempts to separate these signals and recover the input function [99–102], i.e. to deconvolute the signal from the PSF. This technique was brought to digital camera systems to improve sharpness, predicting the defocus of the lens for digital image restoration [103]. Digital Blind Deconvolution has advanced with the improvements in computer science, for instance, successfully measuring and removing motion blur [104], using scene features, such as slanted edges [105], as well as using neural networks [106, 107].

Utilising such techniques to obtain the system MTF has not been studied in great depth due to Blind Deconvolution being an ill-posed problem [108]; it has many solutions for the same goal, with accuracy conditional on the input data signal. Depending upon the approach, Blind Deconvolution works well with astronomy, as there are many point sources for the PSF estimation. For other applications, such as automotive and consumer cameras that contain varied content and a changing PSF across the field of view, Blind Deconvolution would be more problematic.

A comparable method to estimating the system MTF from natural scenes using a Blind Deconvolution style method would be the PSF-CNN approach (cf. § 2.8.4).

2.8.4 Neural Network MTF Estimation

2.8.4.1 PSF-CNN approach

The most successful method to date to obtain camera MTFs from natural scenes utilizes a Convolutional Neural Network (CNN) to estimate the PSF, as proposed by Bauer et al. [15]. The CNN was trained using natural scenes captured using linear camera pipelines, characterized using a PSF panel, i.e. a panel of back-illuminated pinholes. The resulting network estimates the local image sagittal and tangential MTFs, with the input image cropped into 192x192 pixels segments. RAW images were used to avoid compression and system linearisation. This method was designed to take the MTF from the individual colour channels, but the study simplifies the results by only describing the green channel

performance. In this publication, only natural scenes focused on planar subjects were used, minimising blur caused by the depth of field defocus (cf. § 2.9). A Gaussian process regression was applied to the data to smooth and interpolate the noise estimates, obtaining the change in MTF across the field of view. The resulting MTF estimates have high accuracy in relation to the characterized data across several DSLR systems [15] (cf. § 7.3).

Advantages include the MTF being estimated across the entire frame with relatively low computation time, at a few minutes per image [15]. This is not a real-time measure, but it provides MTF data quickly compared to obtaining it under controlled laboratory conditions. A single image can be used for measurement, but multiple images were shown to improve accuracy. The method's drawbacks are that it is not compatible with a standardised method, and importantly, the ground truth is difficult to obtain (needs specialist equipment). Also, as imaging technology advances and camera system performance improve, the CNN may require retraining. This approach has the potential to be developed to obtain many of this project's goals (cf. § 1.2); however, like with all neural networks, it is limited by the training dataset. The CNN is currently trained using several linear pipelines using a *Canon* DSLR with various lenses; there is no information on the accuracy of other camera brand pipelines, as well as the effect of non-linear ISPs and whether an SPD-MTF is possible.

2.8.4.2 Texture-to-MTF Prediction

Bohra et al. introduced the *TextureToMTF* CNN to estimate the e-SFR MTF50 (cf. § 2.6.1) from natural scene patches [16]. This CNN was trained using 959 printed natural scene test charts surrounded by a slanted edge square. Many scene types were used, including text (typed and handwritten), drawing and photographic images. These charts were photographed multiple times using hand-held smartphones, introducing motion blur and a variety of lighting conditions. The authors have made this dataset open access [109] and refer to it as the generalized SFR dataset. The MTF50 from two vertical and two horizontal e-SFR measurements alongside 300 randomly selected patches from the natural scene, sized 48x48 pixels, were used to train the *TextureToMTF* CNN.

This work was conducted to produce a method that provides a system performance metric that would improve no-reference IQMs. Results show that this approach has exceeded the accuracy of state of the art no-reference IQMs [16]. This CNN has also been used to map the MTF50 across a frame, depicting the focus point.

This method works for its intended task to predict and rank image blur for IQMs, but there are some aspects to consider if using such an approach for system performance measurements. The first is to test whether such an approach would estimate the entire MTF and determine the expected accuracy. The second is that smartphone pipelines were used to capture the training dataset; thus, non-linear ISP and compression effects would be present. Sharpening would be applied strongly on the surrounding isolated step-edges, with various levels of local processing occurring on the central image patch. As a result, the MTF50 measurements used in this publication will be overestimated; an MTF50P may be more suitable (cf. § 2.6.1). Finally, utilising nearby isolated step edges to measure the e-SFR would not allow for an SPD measure; however, such an approach can be adjusted to potentially obtain ISO12233 equivalent e-SFRs from natural scenes.

2.9 Natural Scene Derived SFR Considerations

Unlike test charts, natural scenes are not characterised and, as a result, present many obstacles to overcome for a successful automated implementation of the slanted-edge method. Some have already been mentioned, including scene textures, non-uniform illumination and tones, and signal clipping (cf. § 2.5).

High-frequency scene textures act as a form of image noise. Thus, using a high vertical-to-horizontal ROI aspect ratio (cf. § 2.4.1) and implementing a filter such as the filtered tails procedure [83] (cf. § 2.7) reduces its effects on the e-SFR.

Non-uniform ROIs should be avoided or corrected for, as discussed (cf. § 2.5.2). In natural scenes, non-uniformity exceeds simple low-gradients caused by uneven lighting and includes non-uniform tonal changes from natural scene objects. Also, the light side and dark side of the ROI may not have the same non-uniformity. Therefore more complex non-uniform correction may be required than that used in the Imatest software [77].

The dynamic range of scenes surpasses the capturing device in many situations, saturating either the highlights or shadows. As mentioned (cf. § 2.5.3), signal clipping is straightforward to detect so that any clipped edges can be simply removed.

Natural scene step-edges contain varied unknown spatial frequency. Some of these input edges would have to be near-perfect step-edges for the slanted edge method to derive the system's performance (cf. § 2.1.3). Many natural scene edges would not contain maximum

modulation across the spatial frequency bandwidth of the system. Hence, such e-SFRs measured would be a function of both the system performance and the scene content. Some scene edges would maintain higher modulation than the rendering capabilities of the system across the bandwidth. These are suitable edges for estimating the system SFR, but with error, as the input modulation is not corrected for in the measurement. While other scene edges would have lower modulation than the rendition capabilities of the system and would be unsuitable for the SFR. Figure 2.3 depicts the frequency content of such step-edges in relation to an example system MTF. Consequently, this project needs to deselect low-frequency edges (i.e. step-edges that contain low modulation in the high frequencies), leaving the near-perfect and, more likely, *suitable* step-edges that can estimate system performance effectively.

In contrast with two-dimensional test charts, natural scenes have depth. Thus, depth of field needs to be a consideration. As with the PSF-CNN approach (cf. § 2.8.4.1), natural scenes focused on planar subjects can be used to avoid defocused edges. This may not be required, as the out-of-focus edges can also be identified as part of the low-frequency edge detection required to obtain the suitable step-edges.

If the e-SFR is derived from multiple natural scenes, their focus distances may vary. The performance characteristics change with the distance at which an optical system is focused. The extent of this near and far field focus variation is dependent on the optical design. Image file Exchangeable Image File Format (EXIF) metadata may contain the reported focus distance from the lens, but not all systems provide this information, and those that do are not always accurate.

A possible solution to obtaining focus distance is to detect the scene's sharpest edges by measuring their LSF, combined with a depth map of the scene. An accurate distance depth map can easily be created through stereo imaging [110], but utilising two images is not appropriate for this research. Many studies have established methods to obtain single image (monocular) depth estimation [111–115]. Implementing Alhashim and Wonka's proposed method [114] provides a good representation of monocular depth map results, shown in Figure 2.19.

Monocular depth estimation has caveats. First, these monocular depth methodologies contain neural networks to achieve the depth map, and the result depends on their specific training database. Second, the applications reviewed all provide relative distance maps; in



Figure 2.19 Four examples of implementing monocular depth estimation [114].

other words, the map has no distance scale. Thus, to date, there is not a monocular depth estimation suitable for focus distance e-SFR analysis, but it will be a topic to revisit as computer vision technologies improve.

Motion blur is another factor that is avoided when capturing images under controlled conditions. However, motion blur is common when capturing natural scenes, particularly in low light. Any additional blur to the system will degrade the overall performance and is dependent on the direction of the blur in relation to the edge. Theoretically, motion blur perfectly parallel to a step-edge and the pixel array does not degrade performance and reduces noise by removing unwanted scene textures. Whilst motion blur perpendicular to the edge will significantly decrease performance.

Motion blur can be detected, modelled and removed using Inertial Measurement Unit (IMU), an inbuilt system measurement of the angular orientation of a camera in terms of the roll, pitch, and heading angles [116]. This method is most effective when the motion blur is less than 4 pixels. In this project, motion blur is avoided by deselecting blurred edges in the same way as low-frequency and defocused edges. Future research may utilise IMU data to estimate and compensate for the degradation of motion blur on the e-SFR. This would benefit applications with highly characterised camera systems, such as aerial photography and autonomous driving.

2.10 Image Noise

Image noise can significantly corrupt the e-SFR (cf. § 2.5.1); noise is a significant limitation affecting the system performance, especially at low luminance levels. There are many sources of noise, including randomness of photons in a uniform exposure and the resultant photoelectric effect (shot noise), quantization noise and noise from electrical components. Sources of electrical noise include dark current noise, readout noise, reset noise (floor noise), amplification noise and pixel response non-uniformity (PRNU) [6].

The type of noise has a different effect on the overall image. Many noise types change between each frame and are classed as temporal noise. Fixed pattern noise (FPN) refers to noise patterns that do not change from frame to frame.

The primary source of FPN is the dark current non-uniformity; each pixel has a unique dark current across the sensor array, which can increase with heat and exposure time [117]. The PRNU noise is caused by slight variations in the sensitivity of the pixels created by natural variations in the manufacturing process. Therefore, it is not the same as FPN but is constant frame-to-frame, so FPN and PRNU are commonly measured collectively [118].

The sensor design, ISO setting, and shutter speed will impact system noise level. Also, each colour channel may have different amounts of noise due to the relative illumination strength available for each colour channel and depending upon the CFA used [119].

Generally, the ISO settings on digital cameras is a way to change the sensitivity of the system. When increasing the ISO, the sensitivity of the digital sensor is not increased. Instead, the Analogue to Digital Conversion (ADC) gain is amplified, increasing the system noise along with the signal [120].

Four main noise types are commonly focused on when evaluating or simulating a camera system [48, 121]:

1. The photon shot noise, σ_{SHOT} , is caused by the random arrival of photons on the sensor. This random process obeys Poisson statistics in electrons or photons [122], and the magnitude of shot noise increases with illumination. An exposure of q quanta (or electrons) provides shot noise of \sqrt{q} .
2. The read noise, σ_{READ} , is the noise from the on-chip preamplifier caused by a combination of the pixel not resetting to zero and noise associated with the readout

of the sensor array. It has two parts, Gaussian white noise, which decreases as the gain increases, and flicker noise, which is inversely proportional to the input signal affecting the low frequencies. The variance caused by the read noise is approximately equivalent to a distribution of 1 to 3 electrons.

3. Dark current noise, σ_{DARK} , has two parts. The first is shot noise, which follows the Poisson distribution caused by voltage crossover between the pixels in the array. The second is the dark current non-uniformity FPN, discussed earlier, which can be modelled using a Gaussian distribution.
4. Finally, the FPN, σ_{FPN} , is the variation in the signal caused by the discrepancy in pixel sensitivity, the CFA, the gain amplifier and the ADC over the sensor array. Besides the dark current non-uniformity noise, FPN has two other noise types: first, the photon response gain non-uniformity noise, σ_{GAIN} , and secondly, the row-wise fixed pattern noise. Both can be expressed as a Gaussian distribution.

The output signal, V_o , in volts, is expressed as [48]:

$$V_o = (v + \sigma_{SHOT}) \cdot \sigma_{GAIN} + \sigma_{READ} + \sigma_{DARK} + \sigma_{FPN} \quad (2.42)$$

where v is the input signal in volts, and σ is the standard deviation introduced by the various noise types discussed above.

System noise evaluation is standardised in the ISO15739:2017 [4]. This standard allows the total system noise, the temporal noise and the FPN to be measured. A minimum of eight exposures of the uniform diffused target should be taken, with 13% of a system's saturated exposure. These exposures are cropped to 64 x 64 pixels.

The total, temporal and pattern SNRs are calculated by:

$$SNR = \frac{g_{SNR} \cdot L_{SNR}}{\sigma} \quad (2.43)$$

where g_{SNR} is the incremental gain, which is the first derivative of the system's OECF, the L_{SNR} is the luminance equivalent to 13% of system saturation, and σ is the average noise deviations, calculated using the eight images. The total noise, temporal noise, and FPN are calculated using Equations 2.44, 2.45 and 2.46.

$$\sigma_{TOTAL} = \sqrt{\frac{1}{n} \sum_{j=1}^n \sigma^2_{total,j}} \quad (2.44)$$

$$\sigma_{TEMP} = \sqrt{\frac{n}{n-1} \sigma^2_{diff}} \quad (2.45)$$

$$\sigma_{FPN} = \sqrt{\sigma^2_{ave} - \frac{n}{n-1} \cdot \sigma^2_{diff}} \quad (2.46)$$

where n is the number of exposures and j is the exposure number. σ^2_{ave} is the variance in the average of the n exposures, and σ^2_{diff} is the variance between each of the n exposures and the average exposure.

2.11 Summary

This chapter has provided the foundation of knowledge for camera system performance, with particular interest around the slanted-edge method to determine the e-SFR. Topics include the MTF, image formation, and sources of measurement error. The literature is reviewed with respect to adapting the ISO12233 e-SFR for natural scene step-edge inputs.

The slanted-edge method is a robust, well-established standardised methodology to obtain a system's e-SFR. If appropriate step-edges are isolated from natural scenes, the e-SFR can be taken across the field of view, providing a detailed assessment of the system performance.

The ISO12233 requires specific input parameter values to decrease measurement variation. However, these parameter values are restricting when it comes to natural scene step-edge extraction. Therefore, each parameter range should be studied to define suitable ranges that maximise the number of scene edges used when estimating the system e-SFR from natural scenes.

Several factors that affect e-SFR accuracy are controlled in traditional lab-based measurement, but are unavoidable when utilising pictorial natural scenes. These include:

- Step-edges that are not near-perfect, i.e. edges that contain modulation equal to one across the system's bandwidth (scene content) (cf. § 2.1.3 & 2.9)
- Step-edges that contain lower modulation across spatial frequencies than the system capture capabilities, i.e. unsuitable edges to determine the system performance (scene content) (cf. § 2.1.3 & 2.9)
- Defocused edges due to depth of field (lens aperture and focus distance) (cf. § 2.9)
- System noise and scene texture (system engineering and scene content) (cf. § 2.5.1)
- Non-uniformities in illumination and scene tone (scene content) (cf. § 2.5.2)
- Signal clipping (system engineering and exposure settings) (cf. § 2.5.3)
- Various focus distances (system engineering) (cf. § 2.9)
- Motion blur (capturing conditions) (cf. § 2.9)

Some of these are directly researched in this thesis, but others are incorporated as a single deselection process.

Chapter 3 Spatial Frequency Response Derived from Natural Scene Images

The research presented in this chapter explores an automated methodology to obtain a Natural Scene derived Spatial Frequency Response (NS-SFR). The ISO12233 slanted-edge method [5] was adapted to measure e-SFR from step-edges in pictorial natural scenes. The resulting e-SFRs are referred to as NS-SFRs.

In this chapter, the NS-SFR framework is detailed, describing how scene step-edges are detected, isolated, and validated; the rationale for each step is provided. The selected natural scene step-edges are processed with the standard slanted edge algorithm. This chapter concludes with examples of NS-SFR envelopes (range of SFRs) derived from ideal natural scene images, i.e. well-illuminated scenes containing easily isolated step-edges. Results demonstrate that the NS-SFRs are a function of both the system performance and the scene contents.

3.1 Overview of the NS-SFR Framework

This section summarises the proposed NS-SFR framework, with the subsequent sections presenting the main steps of the method in greater depth.

The proposed framework comprises a series of processes to locate, isolate, and validate step-edge ROIs from pictorial natural scenes. These edges are used to measure the NS-SFR through the ISO12233 slanted edge algorithm [5] (cf. § 2.3.2). The framework is either implemented using a single image or a sequence of images (i.e. an image dataset). All images in a dataset must originate from the same camera system, captured with identical optics and lens aperture, since these settings influence the system performance.

Figure 3.1 presents a flowchart of the NS-SFR framework. First, each image from a dataset is selected, rotating images in portrait orientation by 90 degrees to be landscape to ensure that the sensor orientation is constant. Next, the images are linearised using either the Adobe RGB [54] or sRGB [55] linearisation transfer function, as discussed (cf. § 2.3.1). The function used depends on the colour space stated in the image's EXIF metadata. Also, each image is normalised to allow the same computations, irrespective of the image bit depth.

Edges are detected using the Canny edge detector [123] (cf. § 3.2), adapted to separate the vertical and horizontal edge orientations. The resultant horizontal edge binary image is rotated to conform with the edge direction of the vertical counterpart, allowing for the same processing to be applied to both orientations. A proximity filter is applied to both binary images to remove edges within five pixels of a neighbouring edge (cf. § 3.3.1).

Edge locations are then passed through a ROI selection process that crops ROI windows from the linearised image, which are sized appropriately for both the standard e-SFR and natural scenes edge isolation (cf. § 3.3.1). In addition, an edge isolation filter is used to segregate the edge of interest from unwanted artefacts, such as other edges and scene textures. This process is referred to as pixel stretching (cf. § 3.3.2).

The Canny edge detector locates edges based on gradient magnitude, resulting in the detection of wanted step-edges but also other unwanted edge profiles, including stair-case, line, peak and trough edge profiles. The slanted-edge method requires a step-edge; therefore, the isolated edge gradients are examined to select edges with the suitable profile (cf. § 3.3.2). In addition to reviewing the edge profiles, the surrounding areas are also examined to ensure uniform tones on either side of the edge (cf. § 3.4.1). The directions of the edges of interest are measured; any edges with extreme direction change are segmented where possible, or removed (cf. § 3.4.2).

After these isolated natural-scene ROIs have been verified to contain suitable step-edges, they are processed through the traditional slanted-edge algorithm, producing a series (an envelope) of NS-SFRs for each image. *Sfrmat4* [87] is currently utilised for this purpose (cf. § 3.5).

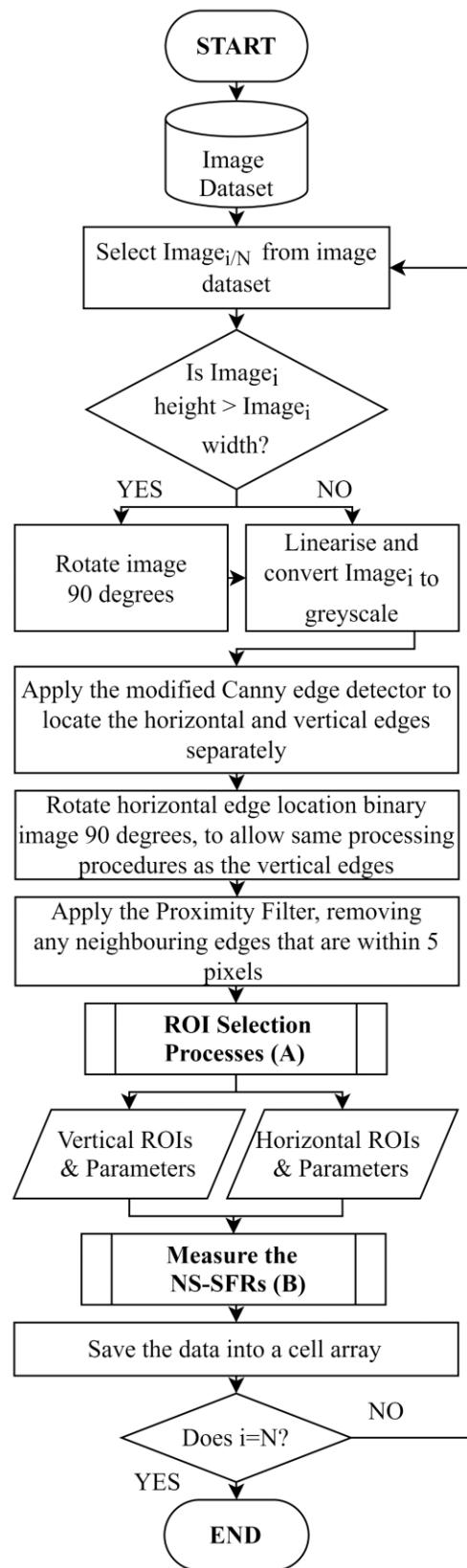


Figure 3.1 Flowchart depicting the core structure of the NS-SFR framework, which extracts, isolates, and validates step-edges from natural scene images before measuring the NS-SFR from a natural scene image dataset. [124].

3.2 Edge Detection

Two automated edge detection workflows were studied for locating and isolating step-edges in pictorial natural scenes, each approaching the task differently. The first employed the Canny edge detector [123, 125], and the second a feature matched filter [126, 127]. This section describes the pilot study that was conducted to compare these two workflows. A small, fixed ROI size (27-pixel width by 41-pixel height) was used in this study, as it was found empirically that it allows natural edges to be isolated with minimal non-uniformity and depth of field changes. The processing required to obtain suitable step-edges is implemented in this pilot study, but more advanced isolation techniques (cf. § 3.3.2) were not.

After testing, the Canny edge detector is chosen for the framework for the reasons detailed below (cf. § 3.2.3).

3.2.1 Canny Edge Detection (Workflow 1)

Canny edge detection is a well documented and implemented edge detection tool, first introduced in 1986 [123] and with many revisions for implementation improvement since its introduction [125, 128–130]. The following steps summarise the established Canny edge detection algorithm [131]:

1. A Gaussian smoothing filter is applied to the image to remove noise.
2. A 2-dimensional (2D) first-order derivative gradient detection filter is applied to the smoothed image. The Sobel operator is used in this workflow, although other common filters include the Roberts and Prewitt operators. The result is x and y orientated gradient ridges.
3. The x and y gradients are then combined to calculate the magnitude.
4. Non-Maxima suppression is applied to obtain the gradient ridge peaks for edge localisation. In other words, this is a skeletonisation process.
5. Hysteresis thresholding is then applied. Two thresholds are used, an upper threshold and a lower threshold. A pixel's gradient above the upper threshold indicates a strong edge and is marked with a one. A pixel gradient below the lower threshold indicates a textured area, and it is removed, i.e. marked with a zero. A pixel gradient is between the two thresholds indicates an area that contains both edge gradients and texture. Thus, this area is accepted if one of the 8-connected gradients surrounding

the pixels contains a strong edge gradient. The output is a binary image indicating the edge locations.

The result is a precise edge localisation that is not computationally heavy. Furthermore, the hysteresis thresholding allows the result to be robust against image noise, ensuring that long edges are not broken down.

For the workflow implementation, the Canny edge detection algorithm was modified. Rather than calculating the magnitude, in step 3, the x and y orientations are kept separate. This adjustment provides horizontal and vertical edge orientation locations in separate binary images. The rationale for the separation is that the e-SFR performance is dependent on edge orientation (cf. § 2.6.2).

The framework uses the Otsu method to determine the hysteresis thresholds [128, 129]. The Otsu method predicts the upper threshold using *Adaptive Particle Swarm Optimization* (APSO), then the lower threshold through constructing a probability model [129]. Applying adaptive thresholds to the Canny edge detector allows edges to be detected more accurately on an image-to-image basis, locating the sharpest edges from the scene whilst selecting fewer false edges and scene texture.

The drawback to using the Canny edge detector is that it detects edges of many gradient profiles, not only the required step-edges. Therefore, once the edges are located, this workflow requires a sequence of processing to validate and select step-edge ROIs.

3.2.2 Matched Filter (Workflow 2)

The feature matched filter used is a normalised 2D cross-correlation filter [126, 127]. This technique locates regions of an image that contain similarity to a desired feature. The desired feature is defined by an input template, which depicts any shape or object. The normalised 2D cross-correlation is summarised as [126, 127]:

1. The cross-correlation of the image with the template is computed.
2. Local sums are calculated through precomputing running sums.
3. These local sums are used to normalise the cross-correlation, forming correlation coefficients.

The normalised 2D cross-correlation, $\gamma(u, v)$, in the spatial domain is defined as [127]:

$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2 \right\}^{0.5}} \quad (3.1)$$

(where $f(x, y)$ represents the image coordinates, $t(u, v)$ the feature template coordinates, $\bar{f}_{u,v}$ is the mean value of the image within the feature region window, and \bar{t} is the mean of the feature).

The 2D cross-correlation was applied using the MATLAB *normxcorr2* function from the Image Processing Toolbox [132]. The 2D cross-correlation was applied in the frequency domain by convolving the Fourier transformed image and feature template, lowering the computation time required for feature detection [127].

This workflow used the matched filter to detect step-edge features. The feature templates are obtained by simulating a series of ramp-edge ROIs, i.e. gradient step-edges, defined using a *tanh* function (cf. § 4.2.1), Equation 4.1. Due to the normalisation of the 2D cross-correlation, the detected image regions are unaffected by the template tonal or contrast values.

The edge angle was adjusted to detect various angled edges. Through testing the matched filter, the angle tolerance was measured to be ± 2.5 degrees. To obtain an angle range of 2.5 to 42.5 degrees, 16 ROI angles are required with an interval of 2.5 degrees. This range is doubled when mirroring the ROI templates to acquire the angle range -42.5 to -2.5 degrees and doubled again to obtain both vertical and horizontal orientations, resulting in 64 matched filter convolution pass-throughs to detect natural scene step-edges.

Only one ROI window size and one *tanh* function were used in this study. The number of targets would need to be expanded to obtain various ROI sizes and edge gradients, doubling the number of convolutions for every new variable.

This workflow targets only step-edges; therefore, no additional processing is required to filter out unsuitable ROIs for the slanted edge method.

3.2.3 Comparison between Edge Detection Workflows

The two workflows approach step-edge detection in opposing directions; both methods have advantages and disadvantages. Workflow 1 (Canny edge detection) takes a top-down approach, selecting edges with an assortment of edge profiles that need to be down-selected to include only step-edges. On the other hand, workflow 2 (matched filter) takes a bottom-up approach; ROIs are gathered and built up through each pass-through of the matched filter, isolating ROIs with pre-determined parameters.

Canny edge detection is computationally faster but requires additional processing time to down-select the edges. The matched filter is computationally heavy but requires no additional processing. Matched filter computation time can be reduced through processing in the frequency domain and utilising acceleration from a *Graphics Processing Unit* (GPU).

The advantage of Workflow 2 is the accuracy of obtaining step-edges from natural scenes. Nonetheless, there was a significant obstacle with the matched filter; it was proven to be too selective. In natural scenes, the probability of obtaining near-perfect step-edges (cf. § 2.1.3), with uniform tones on either side, is low due to many natural scene step-edges containing unwanted artefacts, including image noise, compression artefacts, signal processing artefacts, scene textures, and neighbouring edges. In addition, the modulation content of the edge across the spatial frequencies will vary throughout the scene (cf. § 2.9). The matched filter is not suited to include such a broad range of possibilities and cannot select all appropriate edges from a natural scene image. This is a consequence of a set range of templates, which cannot account for all possible edge features without excessive computation.

A top-down approach was therefore deemed more appropriate for the development of the NS-SFR measurement method. The Canny edge detection workflow is used in the framework despite additional processing requirements. The pilot study showed that, after the deselection process, there were still unsuitable anomalous ROIs being selected. However, this noise is significantly reduced through further analysis and isolation improvements, such as pixel stretching (cf. § 3.3.2). This approach also allows parameter fine-tuning without redesigning the edge detection process, such as introducing variable ROI window crop sizes to best isolate an edge of interest.

3.3 Region of Interest Selection, Isolation and Verification

This section details the processes implemented after obtaining the edge location binary images from the Canny edge detector. First, the ROI windows are cropped from the linearised image, using the detected edge locations (cf. § 3.3.1). Second, the edge of interest within each ROI is isolated using the pixel stretching filter (cf. § 3.3.2). Finally, each ROI passes through a step-edge validation process (cf. § 3.3.3).

3.3.1 Region of Interest Selection

Before cropping the ROI windows, the binary images go through a series of pre-processes. The horizontal and vertical edge locations are compared, any image location with overlapping edge coordinates is removed. The horizontal binary image is rotated 90 degrees so that both edge orientations have the same directional ROIs. This allows all the following processing to be the same for horizontal and vertical edges.

The binary images are then filtered to remove edges in close proximity. Edges that are too close together cannot be used in the slanted edge method due to ESF overlap. In addition, close proximity edges commonly relate to scene areas of high busyness, which is not ideal for isolating step-edges without unwanted scene artifacts (scene texture and other edges).

The proximity filter can be broken down into the following processes; first, a new matrix is created, the same size as the binary image being used, BW_p . The distance between each neighbouring edge location coordinate per image row is calculated. This calculation is implemented by taking the first pair of neighbouring coordinates, x_M and x_N . The threshold T_1 is used to determine whether the edge coordinates are at an appropriate distance. If the distance is greater than the T_1 the two coordinates are accepted, and their coordinates in BW_p are set to one ($BW_p(x_M, y)$ and $BW_p(x_N, y)$); otherwise, they are both set to zero. This process is then repeated for the next pair of neighbouring coordinates in the row, this time x_N becomes the x_M coordinate. If x_M was set to zero in the previous cycle, it will remain zero even though the new set of coordinates may be of appropriate distance. This is because the previous coordinate pair is too close and must remain deselected. In the framework T_1 is set at five pixels, meaning all edges need to be at least five pixels apart.

The threshold of five pixels is set based on the system MTF model (cf. § 2.2.3), derived for the DSLR system used in this project, the *Nikon D800* DSLR. As previously shown in Figure

2.6, the out of phase MTF (minimum performance) has an ESF distribution of 3.7 pixels. Hence, these edges must be at least this distance apart to be isolated without ESF crossover. This modelling provides the NS-SFR framework with a method to determine edges of interest that can be successfully isolated with an ESF mask (cf. § 3.3.2). The number of pixels was rounded up to five to add a margin to account for some further performance degradation from other system components (cf. § 2.2.3). The proximity filter does not limit the edge isolation to ESFs with a distribution of 5 pixels and below, as edges of greater gradients will be isolated from scene areas with fewer neighbouring edges. The threshold should be adjusted based upon the camera system of interest, calculating the ESF distribution using Equation 2.22.

A second threshold, T_2 , is employed to define a minimum pixel distance. This threshold is set at 1 pixel; in other words, if two pixel coordinates are touching, they are likely to be from the same staircase edge. Thus, in such cases, they are both treated as one edge coordinate.

Equation 3.2 details these logical procedures per cycle of the proximity filter:

$$\begin{aligned}
 \text{if } (x_N - x_M) \geq T_1 \quad & BW_P(x_M, y) = \begin{cases} 1 & \text{if } BW_P(x_M, y) \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
 & BW_P(x_N, y) = 1 \\
 \\
 \text{else if } (x_N - x_M) = T_2 \quad & BW_P(x_M, y) = \begin{cases} 1 & \text{if } BW_P(x_M, y) \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
 & BW_P(x_N, y) = \begin{cases} 1 & \text{if } BW_P(x_M, y) \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
 \\
 \text{otherwise} \quad & BW_P(x_M, y) = 0 \\
 & BW_P(x_N, y) = 0
 \end{aligned} \tag{3.2}$$

After the proximity filter is applied to both binary images, the height of the selected edges is measured. Any edge found to be shorter than 20 pixels is deselected as it is too short for the purpose of NS-SFR measurements (cf. § 4.4).

At this stage, the binary images are ready to be used to crop the edges into ROIs. The aim is to automatically select ROIs and adjust the window size to best suit each edge. Therefore, the coordinates for the crops are determined through the MATLAB bounding box in the *regionprops* function [133] from the Image Processing Toolbox [132]. This function provides top left corner coordinates with the height and width to give each binary object

(edge) a bounding box. However, these coordinates provide a boundary that presses against the edge width. Thus, all the coordinates are expanded by ten pixels, five pixels on either side of the edge. This is based on the camera system MTF model and ensures that the ROI boundaries do not hinder the ESF distribution. In addition, any widths smaller than twenty pixels are expanded to meet this minimum threshold, minimising the error in the NS-SFRs (cf. § 4.4).

When cropping edges from natural scenes, tall edges are more likely to contain unwanted artifacts, such as depth of field changes across the edge, intersecting edges and illumination non-uniformity. Thus, small ROIs are desired from natural scenes, however, too small ROIs have been shown to introduce bias to the e-SFR (cf. § 2.4.1); a height of 128 pixels provides a good balance (cf. § 4.3), as well as being a commonly used benchmarking ROI size [22, 63, 134]. Therefore, the height of the box boundaries is also adjusted; with the goal height of 128 pixels, any edge greater than this height is broken down into smaller boundaries. The goal height is prioritised before extracting smaller boundary heights. These new boundaries are governed by previously set thresholds, including the minimum edge height and boundary width.

Once boundary coordinates are established, they are used to crop the ROI windows from the linearised image. The image is rotated 90 degrees for the horizontal edge crops to match the previously rotated binary image.

Figure 3.2 illustrates the processes in the framework to isolate step-edges. This flowchart also details edge masking before edge isolation (cf. § 3.3.2) and step-edge validation (cf. § 3.3.2).

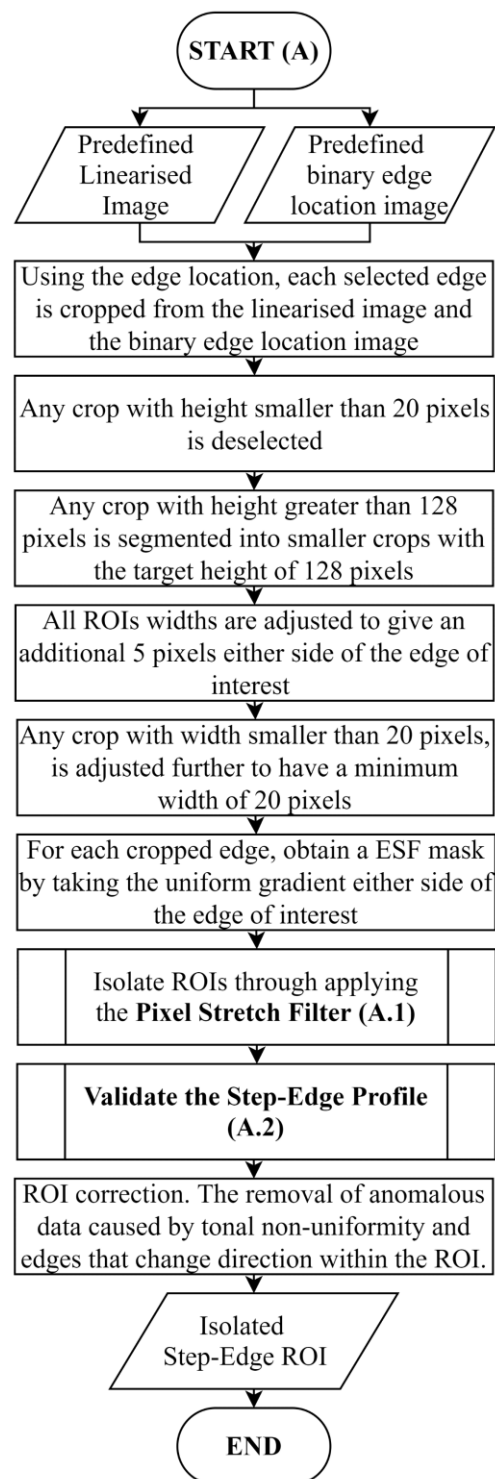


Figure 3.2 The ROI selection process, adjusting the box boundaries for cropping the ROIs, measuring the ESF mask and applying the pixel stretching (cf. § 3.3.2), then validating that the selected ROIs contain the required step-edge profile (cf. § 3.3.2). The position in the framework is shown in Figure 3.1, marked with *A*. [124].

3.3.2 Edge Isolation – Pixel Stretching

There are image structures other than the desired edge of interest in many ROIs extracted from pictorial natural scenes; it is rare to find a scene step-edge with uniform tones on both sides of the ROI. Therefore, once the potential ROIs are cropped, the edge of interest is isolated, ensuring any unwanted scene content is removed from the ROI. This isolation is achieved by developing the pixel stretching filter, influenced by the filtered tails procedure [83] (cf. § 2.7.1), a filter that blurs either side of an edge to reduce noise error.

Before applying pixel stretching, the ESF mask for each ROI is first obtained. The ESF mask is the area of the ROI where the edge and its distribution are located; it is created by finding where the gradient on either side of the edge location becomes uniform. A threshold is set to account for the image noise floor, a static normalised pixel value of 0.02, anything under this value is considered uniform. This value was empirically determined to detect image noise and low contrast scene content. Ideally, it should be adaptable to the individual ROIs, accounting for image noise and scene texture. Next, a dilation of three pixels is applied to expand the ESF mask border to ensure that the ESF is within this masked area. The ESF mask is illustrated in Figure 3.3. The ROI crop and the corresponding ESF mask are used as the input to the pixel stretching. This procedure is depicted in the flowchart in Figure 3.4



Figure 3.3 Depiction of the ESF mask [76].

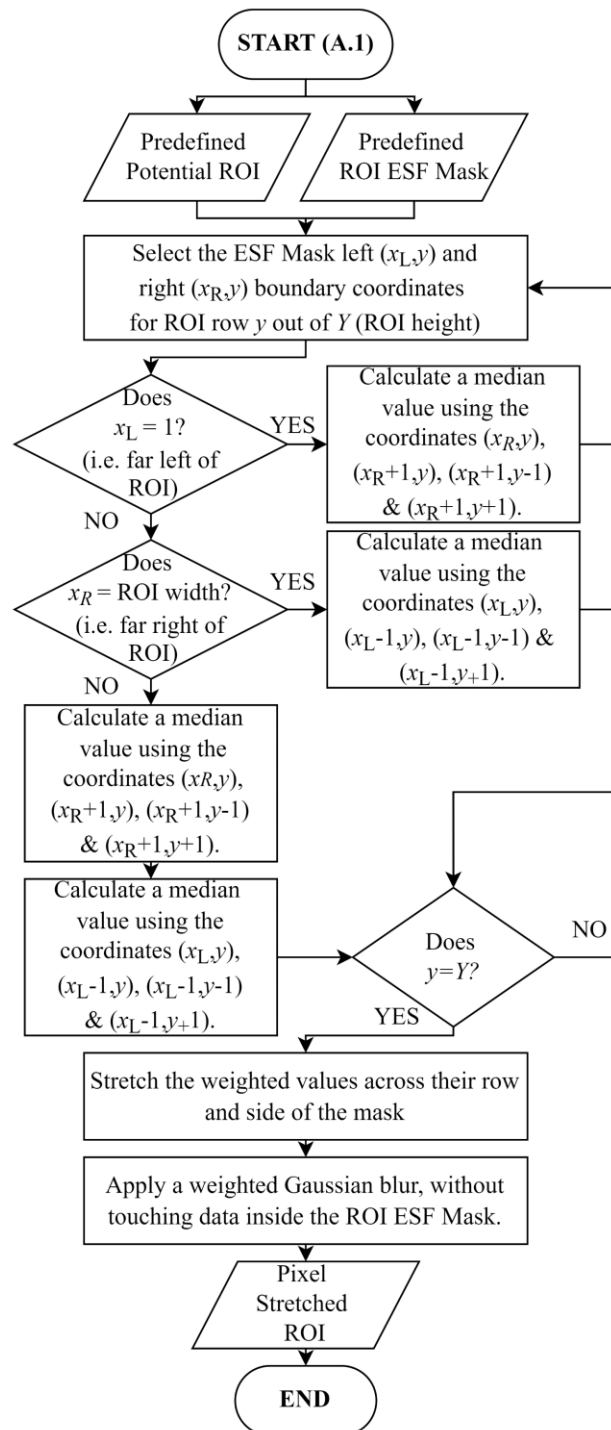


Figure 3.4 Flowchart highlighting the steps for the pixel stretching filter. The position in the framework is shown in Figure 3.2, marked with *A.1*. [124].

The mask boundary coordinates to the left (x_L, y) and right (x_R, y) are located for each pixel row in the ROIs. Two T -shaped median values are measured from neighbouring pixel values. The left value is measured using the coordinates (x_L, y) , (x_L-1, y) , $(x_L-1, y-1)$ and $(x_L-1, y+1)$. Whilst the right median is calculated from (x_R, y) , (x_R+1, y) , $(x_R+1, y-1)$ and $(x_R+1, y+1)$. These median values are then used to fill the entire row of the ROI without touching the ESF within the masked area, as illustrated in Figure 3.5. Only one side is stretched if the ESF mask is at the ROI boundary.

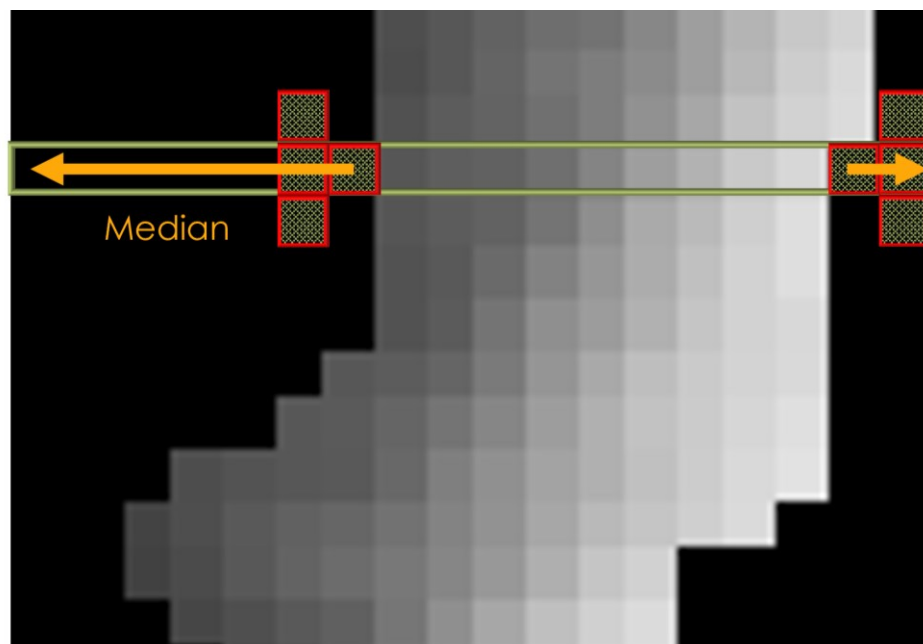


Figure 3.5 Illustration of the T -shaped median coordinates and how they were used to fill each pixel row, i.e. stretching the median pixel value. [76].

This pixel stretching filter can introduce a streaking artifact in the ROIs, caused by changes in the median value. Although sampling multiple rows by the T -shape median does help to reduce this artifact, under high image noise, streaking may still occur. Therefore, a weighted Gaussian blur is applied to reduce this artifact. The filter is weighted one in the opposite corners of the ROI furthest from the edge and is reduced to zero as it approaches the ESF mask boundaries, preserving the edge of interest. The weight values from zero to one follow a square power function, meaning the rate of change increases as it approaches one.

Figure 3.6 illustrates the pixel stretching filter on two simulated edges, a noiseless ROI and a ROI containing a high noise level (SNR 4). As intended, the filter helps to isolate the edge of interest. The streaking artifact is present in the ROI with high noise, and the blur from the weighted Gaussian is demonstrated.

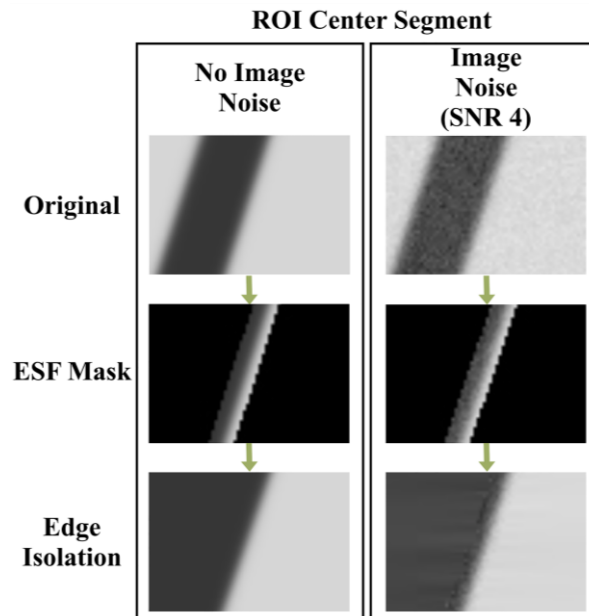


Figure 3.6 An example of edge isolation using the pixel stretching on simulated ROIs, one noiseless and the other with a noise level of SNR 4. [76].

After testing the pixel stretching on simulated step-edges created using a \tanh function, Equation 4.1 (cf. § 4.2.3), at various levels of simulated image noise (cf. § 2.10), it was observed that, as long as the ESF distribution is not affected, no error is introduced to the e-SFR compared to the noiseless edge input.

The pixel stretching was shown not only to improve edge isolation, increasing the number of edges that are isolated from a natural scene image at appropriate ROI window sizes, but also to reduce the corruptive nature of noise (cf. § 3.3.2.1) and non-uniformity (cf. § 3.3.2.2) on the e-SFR.

3.3.2.1 Noise Reduction

Image noise introduces positive bias to the high frequencies of e-SFRs [39, 56, 63] (cf. § 2.5.1). The filtered tails procedure [83] (cf. § 2.7.1) is a filter that blurs either side of a noisy step-edge, implemented to reduce this bias effectively. As this procedure inspired the development of pixel stretching, they share the same noise reduction benefits.

Figure 3.7 demonstrates the noise reduction from the pixel stretching filter. Five \tanh ROIs were simulated with different noise levels. These noise levels were simulated using Poisson and Gaussian distributions to model shot and read noise (cf. § 2.10 & 4.2.3). The SNR was

measured using 13% of the system saturation (cf. § 2.10), Equation 2.43. The e-SFRs were measured and plotted without and then with the implementation of pixel stretching.

The results are clear; pixel stretching reduces high-frequency noise corruption. This SFR noise reduction is valuable when extracting edges from natural scenes. Unlike well-illuminated test charts, natural scene images can be captured under various illumination conditions, requiring different system gain settings resulting in different system noise levels. In addition, high-frequency scene textures are also a form of noise that can introduce similar positive bias and be reduced with the pixel stretching filter.

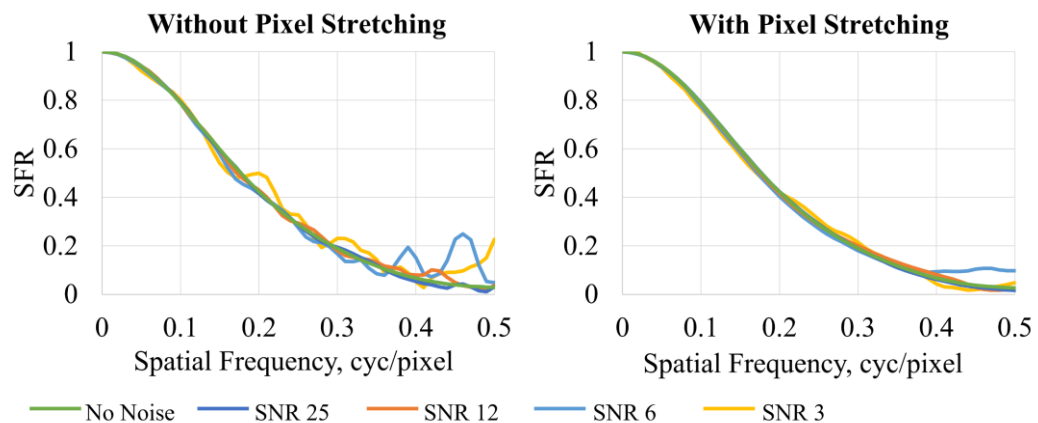


Figure 3.7 These two graphs illustrate the effect that the pixel stretching edge isolation has on ROI with high noise, SNRs 25, 12, 6 and 3. On the left are the resultant e-SFRs from these ROIs without pixel stretching, and on the right are e-SFRs from the same ROIs but applying pixel stretching [124].

3.3.2.2 Non-uniformity Reduction

Non-uniform illumination is problematic when extracting step-edges from natural scenes [77] (cf. § 2.5.2). Due to the pixel stretching filter masking the ESF and filling the ROI with an average value on either side of the edge, much of the non-uniformity is diminished. However, it is not eliminated since non-uniformity is maintained within the ESF mask.

Figure 3.8 demonstrates the effects of pixel stretching on simulated non-uniform ROIs. The non-uniform illumination was simulated by creating low-frequency gradients (step-edge containing no modulation in high-frequencies) at five angles, 0° to 180° . These gradients ranged in pixel values from 0 to 43 (8-bit) and were added to the simulated ROIs.

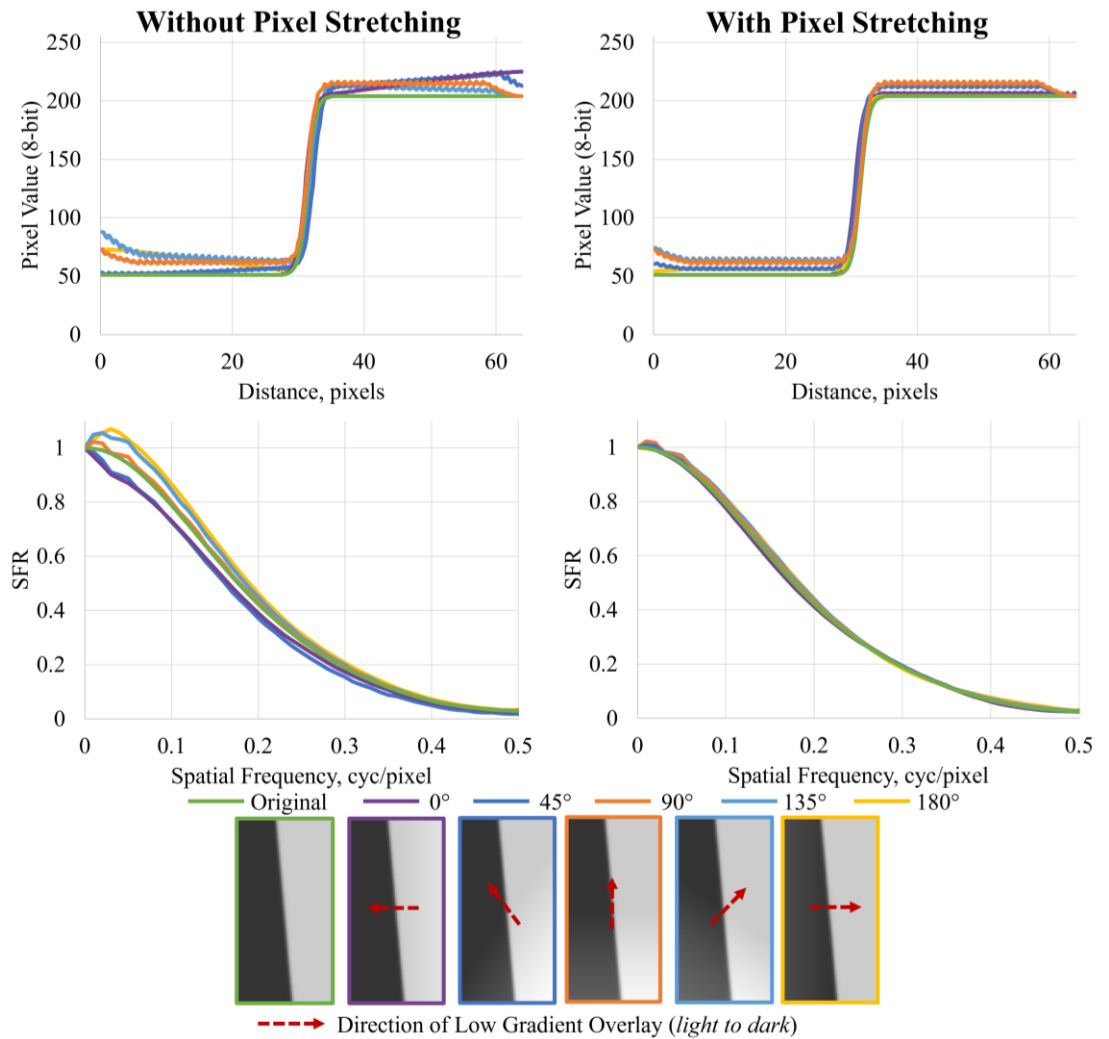


Figure 3.8 A demonstration of how the pixel stretching edge isolation reduces the effects of non-uniform illumination in the ISO12233 e-SFR. The left shows the ESF and SFR from ROIs with a low-gradient overlay. On the right are the same ROIs but using pixel stretching [124].

The simulated data shows how the pixel stretching reduces the tone changes on either side of the ESF. In turn, it stabilises the resultant e-SFR measure. Although the reduction in error from illumination non-uniformity is significant and allows such ROIs to be used, the pixel stretching filter does not entirely remove the error introduced.

3.3.2.3 False Sharpening Artifacts

When applying pixel stretching to natural scene ROIs, it was observed that it would amplify a false sharpening artifact. False sharpening refers to tonal values on either side of the edge, creating a lobe in the ESF, imitating the effects of sharpening algorithms on edges (cf. § 2.4.2). Such ROIs result in NS-SFR overestimation. Note that the pixel stretching does not

introduce the artifact but amplifies it; the scene tonality and illumination around the edge are the most common causes.

Three natural scene ROI examples are shown in Figure 3.9. First, ROI a) illustrates a building corner, where the light and shadow create a highlight and a shadow on either side of the edge. Next, ROI b) illustrates a curved corner, where the diffused reflection around the edge has caused a highlight increase. Finally, ROI c) provides an example of chromatic aberration also causing this false sharpening artifact.

This artifact can be problematic when estimating the system e-SFR (cf. § 5.1). On average, 39% of the NS-SFRs gathered from the DSLR 1 and 2 datasets (cf. § 5.3.1) are above the expected ISO12233 e-SFR envelope (measured between 0 and 75% of the Nyquist frequency, to remove frequencies biased from noise). However, when using an extensive dataset of images, these types of ROI are deselected if the LSF is too wide (cf. § 5.1.3) or averaged out during the system e-SFR estimation (cf. 5.1.4).

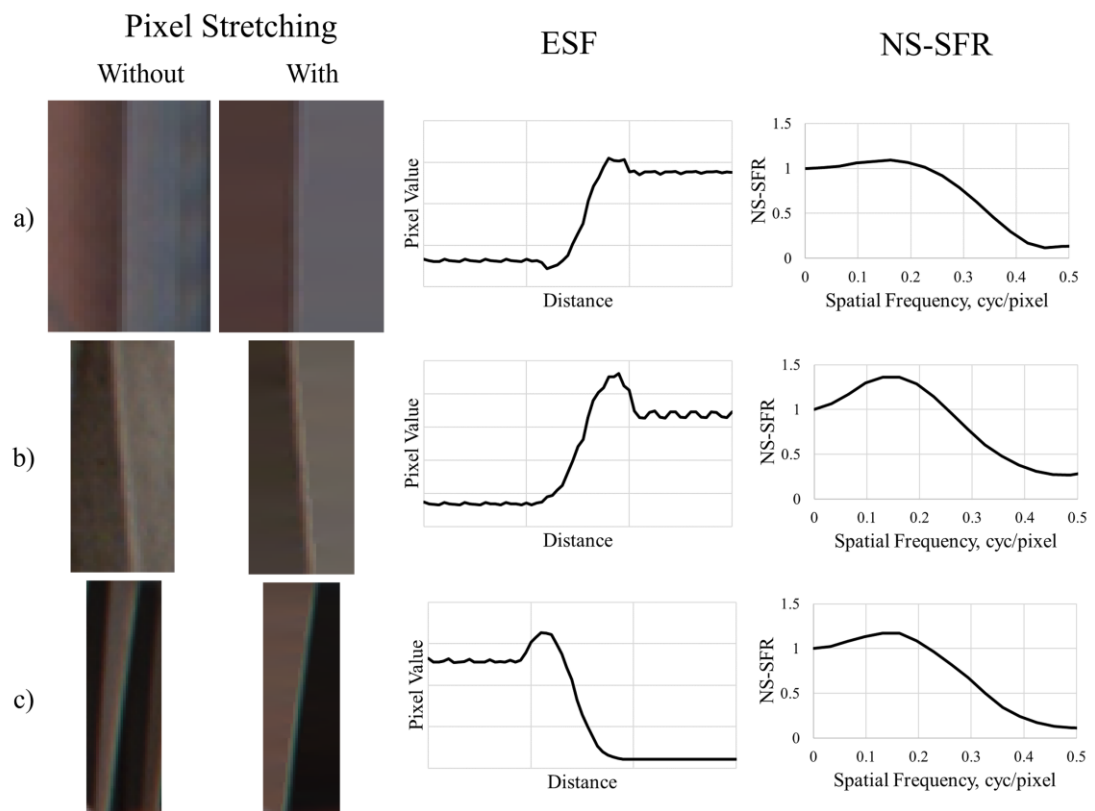


Figure 3.9 Three examples of false sharpening occurring in natural scene ROIs. Each example is illustrated without and with pixel stretching, and the pixel stretched ESFs and NS-SFRs are plotted.

3.3.3 Step-Edge Validation

As discussed (cf. § 3.2), the Canny edge detector locates edges with various edge profiles. Therefore, the framework must deselect edges that do not have the suitable step-edge profile for the slanted-edge method. To achieve this, after the ROIs are cropped and the edge of interest isolated, they are placed through a step-edge validation process.

The principle of the step-edge validation method is straightforward; if there is a single increase or decrease in the edge gradient above the noise floor, then a step-edge is present. This logic is illustrated in Figure 3.10. The method flowchart is outlined in Figure 3.11.

This method is implemented by taking each row from a potential ROI, measuring its gradient. If the gradient meets the requirements, that row is marked as a confirmed step-edge. Any unwanted edge profiles are detected, as shown in Figure 3.10. As with creating the edge mask (cf. § 3.3.2), a static normalised pixel value of 0.02 is used as the gradient threshold to define the uniform values, taking into consideration the noise floor.

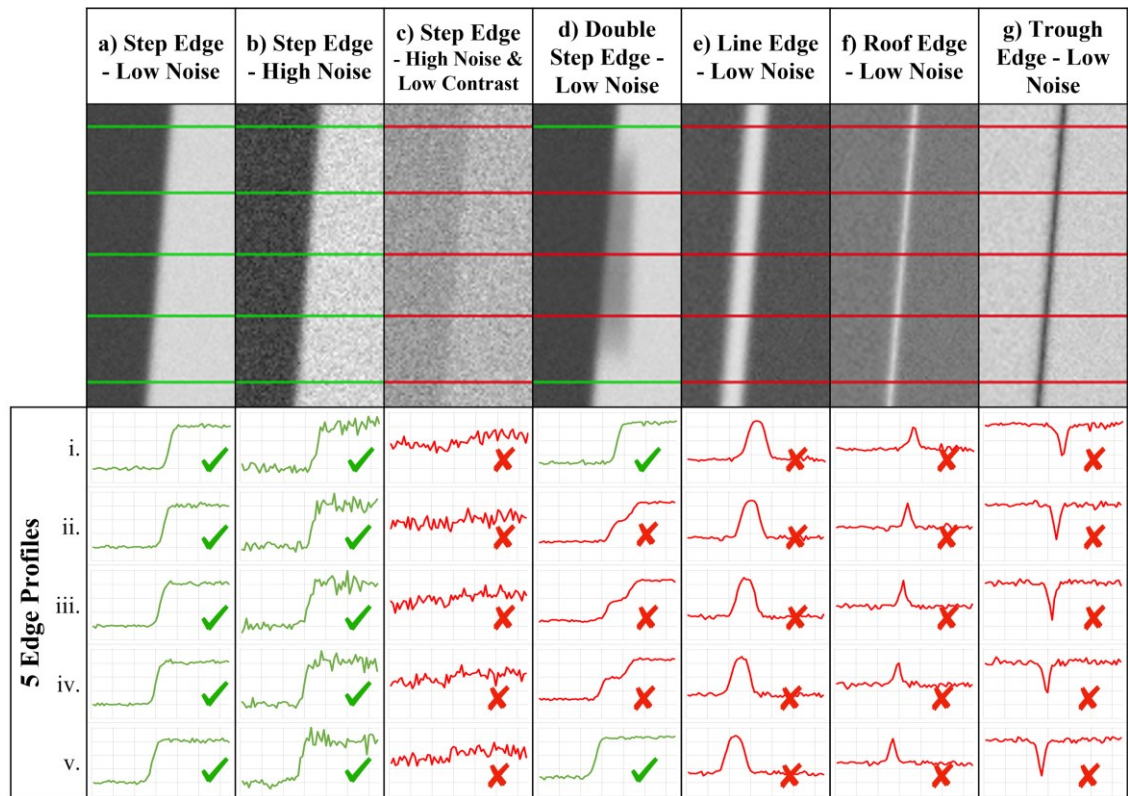


Figure 3.10 These seven ROIs contain various edge profiles, including step edges, staircase edge, line edge, roof edge, and trough edge. This diagram describes how the measured gradient can successfully determine useable step edge ROIs [76, 124].

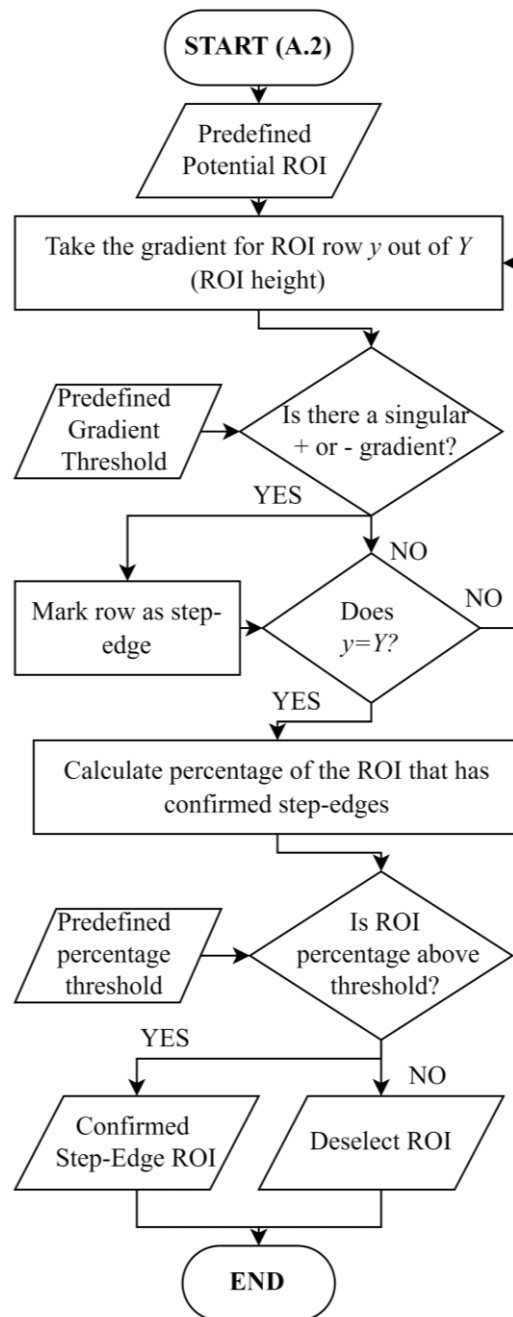


Figure 3.11 Step-edge validation flowchart describing the process to select only step-edge ROIs. This process is marked with *A.2* in the framework in Figure 3.2 [124].

Figure 3.12 further demonstrates how the method reacts with image noise at high and low edge contrast. This figure plots a ROI in three-dimensional (3D) space. Without noise, the single positive gradient is apparent. The gradient is still detectable for the same edge with high image noise. Gradients in the uniform areas of the noisy ROI are highlighted as the noise exceeds the gradient threshold. The edge contrast must be above the noise floor, otherwise the step-edge is not detected.

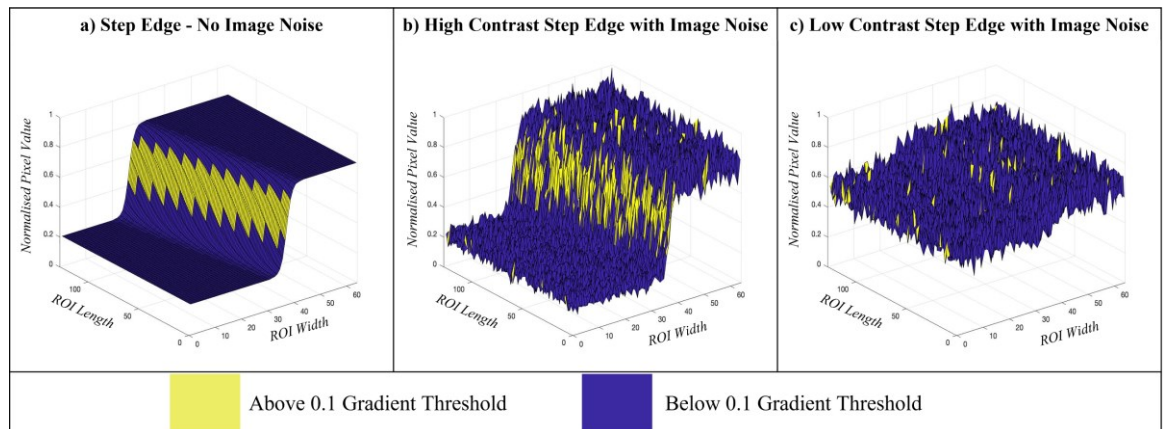


Figure 3.12 Depiction of the step-edge verification process. 0.1 is the gradient threshold in this example, any gradient above this threshold is marked.

A suitable gradient threshold is important. If the threshold is too low, the method will be sensitive to noise. On the other hand, low contrast scene content is missed if it is too high, potentially validating unwanted edge profiles.

Once the ROI row gradients are measured, the percentage of them with a confirmed step-edge is computed. If this percentage exceeds the set threshold, the ROI is selected as containing a step-edge. The percentage threshold is set at 50% of the ROI.

These thresholds are both constant and were determined using an empirical examination of NS-SFRs taken from the DSLR 1 image dataset (cf. § 5.3.1). The gradient and percentage thresholds are set in conjunction. The examination determined the best balance between the two thresholds to obtain robust NS-SFRs, minimising anomalous data.

The gradient threshold is set to be sensitive to low contrast edges, allowing the detection of scene content; however, this will also include system noise. As noise is random, the rows labelled as non-step-edges are randomly distributed through the ROI, whilst non-step-edge rows due to scene content are in concentrated areas of the ROI. The percentage threshold provides a counterbalance to the sensitive gradient threshold. Low contrast scene content is detected, and the ROI removed, whilst this threshold allows ROIs with some noise to pass. ROIs with noise exceeding 50% of the rows is not ideal and are deselected.

The use of a single camera system means that the thresholds are tailored specifically for this DSLR system. Ideally, rather than using static thresholds, further work should determine an adaptive threshold based on the noise floor to better match each natural scene ROI (cf. § 7.2.2).

3.4 Removal of Anomalous ROIs

Although the framework has thus far selected and then validated step-edges from the natural scene data, ROI errors and anomalies are still problematic. After edge verification, several further stages are implemented to ensure that such ROIs are removed before inputting them into the slanted-edge algorithm.

3.4.1 Non-Uniformity Streaking Detection

Pixel stretching helps reduce the effects of low-gradient non-uniformity (cf. § 3.3.2.2). However, it may cause a streaking artifact with high noise or unwanted scene content intersecting the edge of interest (cf. § 3.3.2). This artifact can create areas of non-uniformity under extreme circumstances; selecting a mean value that is too high or too low compared to the surrounding pixels introduces heavy streaking. ROI b) in Figure 3.13 is a simulated example of this.

To detect such areas, each ROI has two vertical gradients measured from either side of the ESF mask boundary. These gradients are demonstrated in Figure 3.13. The framework flags the ROI rows where the gradient increases above the gradient threshold, once again set at 0.02 (cf. § 3.3.2 & 3.3.3).

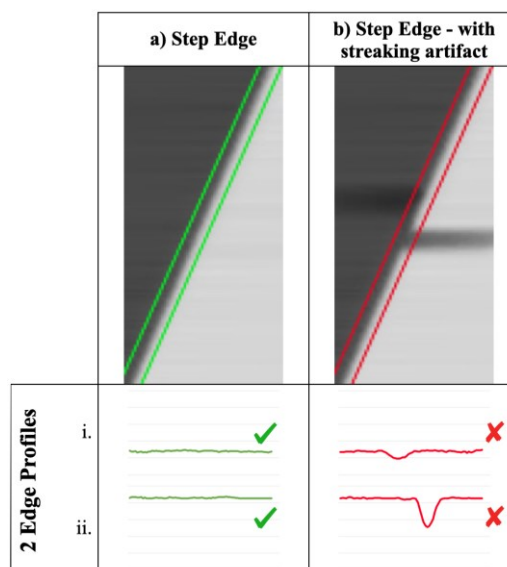


Figure 3.13 Simulated demonstration of the two vertical gradients following the ESF mask boundary. These gradients are used to detect areas of non-uniformity in the isolated ROIs.

3.4.2 Multidirectional Edges Detection

Other than the horizontal and vertical orientation separation (cf. § 3.2.1), the direction of the edges have not yet been examined. Although every ROI contains an edge in one orientation (horizontal or vertical), the edge may change direction. The direction of each edge is therefore examined in the x and y space. Ideally, both dimensions should either have a positive or negative increase throughout the ROI. If the edge has a sudden change in direction, the ROI row is flagged.

Figure 3.14 demonstrates a simulated edge with a) two changes in the x -direction and b) two changes in the y -direction. When there is an x -direction change, the ROI is easily segmented at the flagged ROI row. However, areas of the ROI with a change in the y -direction are unwanted and are flagged.

A slight change in edge direction in the form of curvature is permitted. Curved edges are corrected for in the *sformat4* slanted-edge algorithm [87] (cf. § 3.5).

Once the position of non-uniform areas and changes in the edge direction are detected, the ROI is segmented into new smaller ROIs, removing the rows containing non-uniformity and y -direction changes. The height of the segmented ROIs must be greater than the minimum threshold (20 pixels (cf. § 3.3.1)).

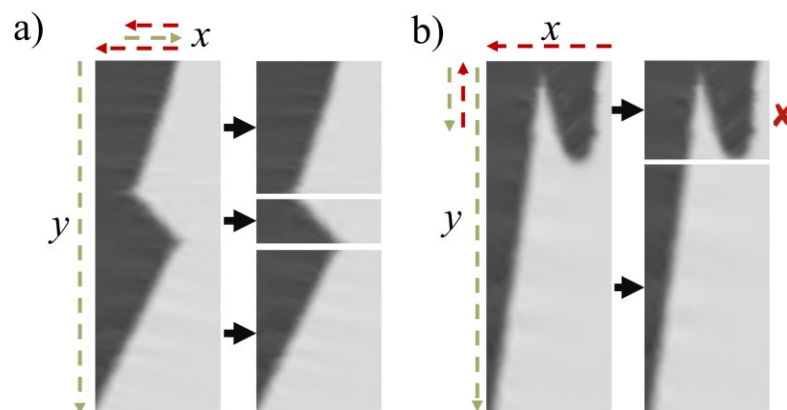


Figure 3.14 Simulated illustration of the detection and segmentation of unwanted changes in the a) x and b) y edge directions.

3.5 The NS-SFR Measure

So far, the framework has extracted, isolated, and verified step-edges from natural scenes. The next stage is measuring the NS-SFRs. This measurement is shown in the Figure 3.15 flowchart.

The SFR calculation is achieved using the ISO12233:2017 e-SFR slanted-edge algorithm [5]. Rather than using linear edge fitting, as described in the standard, a third-order polynomial fitting is used to allow curved edges to be used without introducing error (cf. § 2.7.1) [86, 135]. Burns' established *sfrmat4* MATLAB code [87] is used for this purpose.

Minor adjustments were made to the *sfrmat4* code, which have not impacted the slanted-edge algorithm itself. The adjustments allow edge angle and contrast to be recorded, as well as indicate whether clipping is present; the original *sfrmat4* code was already calculating these three analytical measurements, but they were not output.

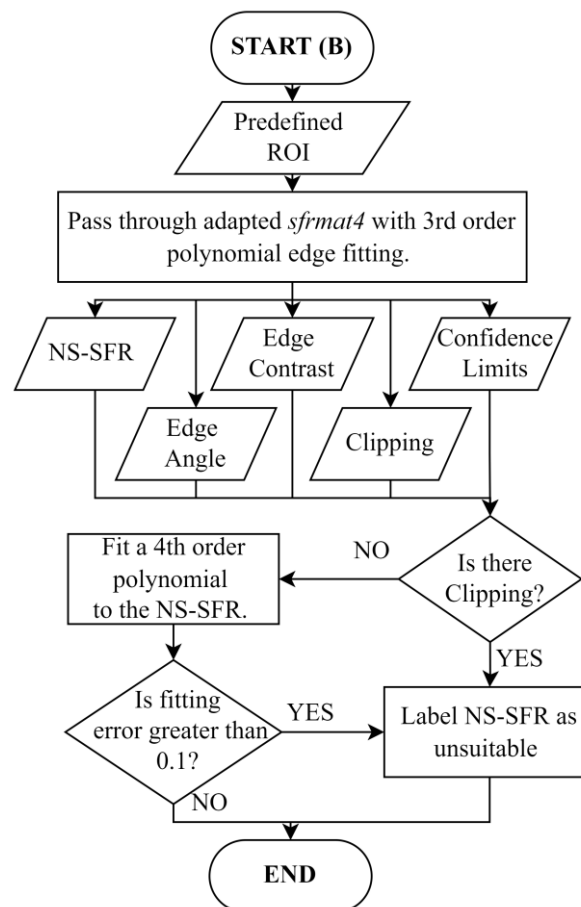


Figure 3.15 Flowchart depicting measuring the NS-SFR from the validated ROIs. This process is marked with ‘B’ in the framework in Figure 3.1 [124].

Unlike test charts, natural scene edges are unlikely to be neutral, resulting in higher luminance contrast in one colour channel; for example, an edge between a wall and the sky will have a predominant blue channel, with low contrast red and green channels. Hence, an NS-SFR cannot always be measured for each channel. Thus, only the luminance values of the ROI are used in this project to simplify and mitigate this issue. This is achieved by the following conversion [136, 137]:

$$L = 0.299 R + 0.587 G + 0.114 B \quad (3.3)$$

where L is luminance, and R , G , B refer to the pixel responses of the red, green, and blue colour channels.

Although extensive ROI isolation and verification is implemented in the framework, unsuitable ROIs still can make it through to this stage. For example, errors caused by incorrect edge location due to the edge having a curvature that cannot be characterised with a third order polynomial, such as a complex ‘S’-shaped edge. These ROIs are identified before an error can occur in the *sfrmat4* code, and the ROIs are deselected.

Once the NS-SFRs in an image are measured, they undergo a process to filter out results with high noise. An e-SFR should theoretically follow half of a bell-shaped curve; however, NS-SFRs can contain noisy scene data, drastically distorting this curve. Such measurements are removed to reduce bias in the results. This is achieved by fitting a fourth order polynomial to each NS-SFR curve. If the absolute error between the data and fitted function exceeds 0.1 modulation, the data is deselected. This modulation threshold was chosen empirically based on DSLR and smartphone NS-SFR data (cf. § 5.3.1). It provides enough tolerance to permit NS-SFRs with sharpening lobes and bias from high-frequency image noise, but is stringent enough to remove the anomalous NS-SFRs that do not conform to the correctly shaped e-SFR curve.

3.6 Initial Testing: The Image NS-SFR Envelope

After the NS-SFR framework was established, tests were conducted to i) ensure correct ROIs were being extracted, ii) observe any correlation between the natural scene and NS-SFRs, and iii) find the relation to the equivalent ISO12233 test chart measure. The DSLR 1 image

dataset (cf. § 5.3.1) was used in this testing; the DSLR 1 camera system was characterised using the established ISO12233 slanted-edge method.

A pilot study was first carried out to determine if the NS-SFR framework can isolate straightforward step-edges before more complex natural scene edges. The test chart image used to characterise the system was passed through the framework. Figure 3.16 plots the vertical e-SFR envelopes, devised from both manually selecting the ROIs, as per the ISO12233, and the framework isolation. SFR envelopes were defined through three curves, the 5th and 95th percentiles of the data, and a weighted mean. These two percentiles define the statistical position of 90% of the measured e-SFR performance, which was selected to remove the outliers. The weights used for the mean were based on the radial distance from which the ROI is located in the frame. The centre ROIs were weighted by 1.00, part-way was weighted 0.75 and corners 0.50. These weights are the same default zone weight values that Imatest use in their MTF evaluation software [74]. Comparing these two envelopes show that the framework isolates the correct step-edges and forms a matching e-SFR envelope. The negligible variation between the two envelopes is due to the bias introduced by the pixel stretching technique, as previously discussed (cf. § 3.3.2). Also, ROI heights of 128 pixels and smaller were isolated when using the framework (cf. § 3.3.1), rather than taller ROIs from a manual selection process, introducing variation.

The NS-SFR framework was applied to a small dataset of 30 natural scene images, captured with the DSLR 1 system (cf. § 5.3.1). These 30 images were intentionally selected to provide

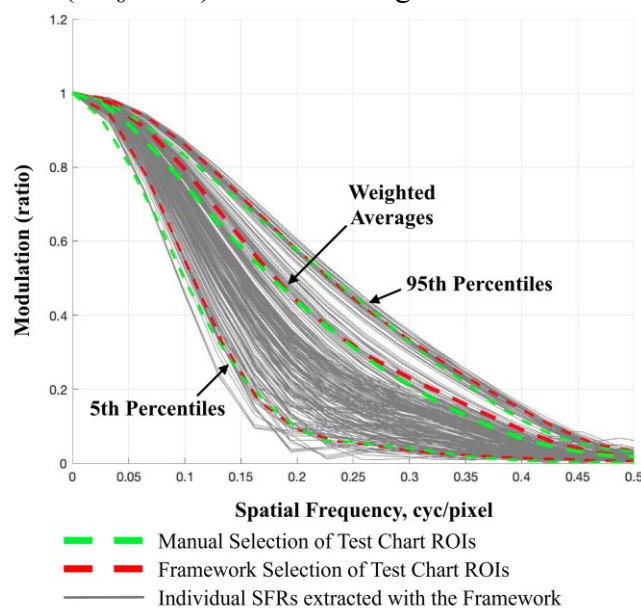


Figure 3.16 The e-SFR envelope compares the manual isolation of test chart edges with those extracted using the NS-SFR framework [76].

optimal conditions for the NS-SFR framework. They are captures of well-illuminated scenes, with minimal system noise, containing many step-edges and captured with a large depth of field.

Figure 3.17 illustrates the vertical NS-SFR envelope for four of these scenes compared to the equivalent envelope devised from the standard method. These envelopes were defined using the same three curves as Figure 3.16, the weighted mean, the 5th and 95th percentiles. This data provided important observations of NS-SFR characteristics, indicating that the NS-SFRs are a function of both the system performance and scene content.

The first observation is that the NS-SFR envelope is highly dependent on the location of the ROIs extracted from the field of view. This variation is because optical performance changes across the frame (cf. § 2.4.3). Thus, when most ROIs are isolated from the frame corners, such as in Image 1, the NS-SFR envelope is positioned towards the lower performance side of the test chart. In contrast, if the majority of ROIs are extracted from the frame centre, such as in Images 2 and 3, the NS-SFR envelope is positioned in the higher performance region of the test chart envelope. When the ROIs are evenly distributed across the frame, such as in Image 4, the NS-SFR and chart e-SFR envelopes match well.

There is an increase in high-frequencies in the NS-SFR of all four images. Usually, this is an effect of image noise (cf. § 2.5.1). Noise, however, is not an issue because these scenes are captured with low ISO gain settings. This observation confirms that the high-frequency scene texture acts as noise, increasing the NS-SFR in the high-frequencies. As mentioned above, Image 1 provides a low-performance NS-SFR envelope, although rather than decreasing to zero as frequency increases, the 95th percentile forms a plateau at 0.2 modulation, a noise floor of the increased frequencies.

Some NS-SFRs are over- and others are under-estimated compared to test chart e-SFRs. The underestimated SFRs are due to the scene content; many natural scene step-edges do not contain higher modulation than the system capture capabilities through the frequency bandwidth, i.e. *suitable* step-edges for a system e-SFR estimation (cf. § 2.9). Thus, after the system degrades the signal, the NS-SFR are lower than the system's performance, shown in Figure 2.3. Overestimated NS-SFRs are because anomalous edge profiles pass through the framework due to the ROI content or false sharpening, as previously discussed (cf. § 3.3.2.3).

Figure 3.18 shows the mean weighted average from all 30 scene images. The result is a close approximation of the average test chart e-SFR envelope. Less suitable scenes are expected to introduce higher variation, due to high image noise, out of focus edges, shallow depth of field and a range of focus distances (cf. § 2.9). Therefore, the next step to estimate the system e-SFR with adequate accuracy from NS-SFRs is to analyse and isolate the data with the least variation introduced by scene and system.

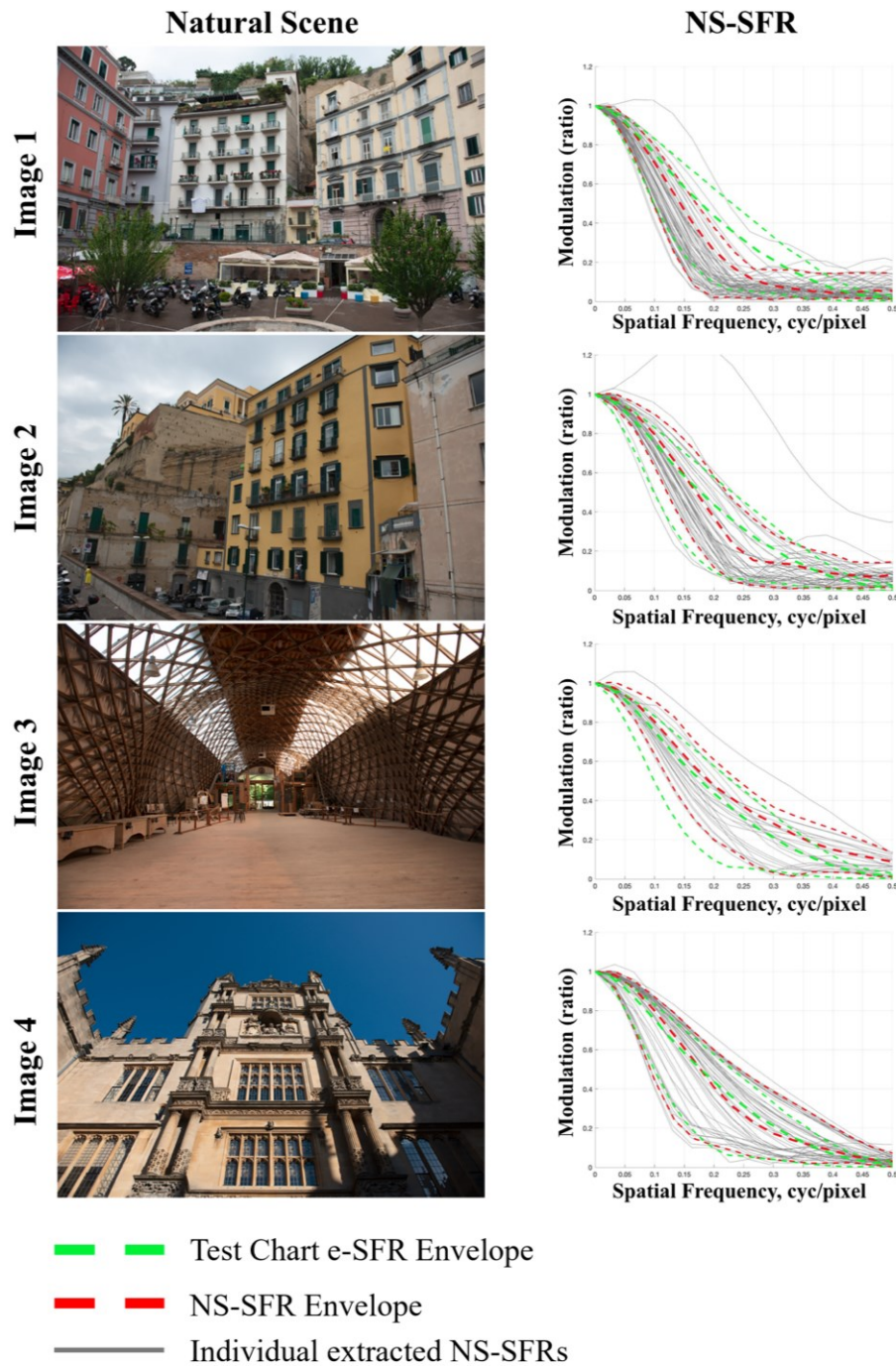


Figure 3.17 The NS-SFR envelope from four natural scenes compared to the equivalent test chart e-SFR envelope. The envelopes are outlined by the 5th and 95th percentiles of the data and a weighted mean [76].

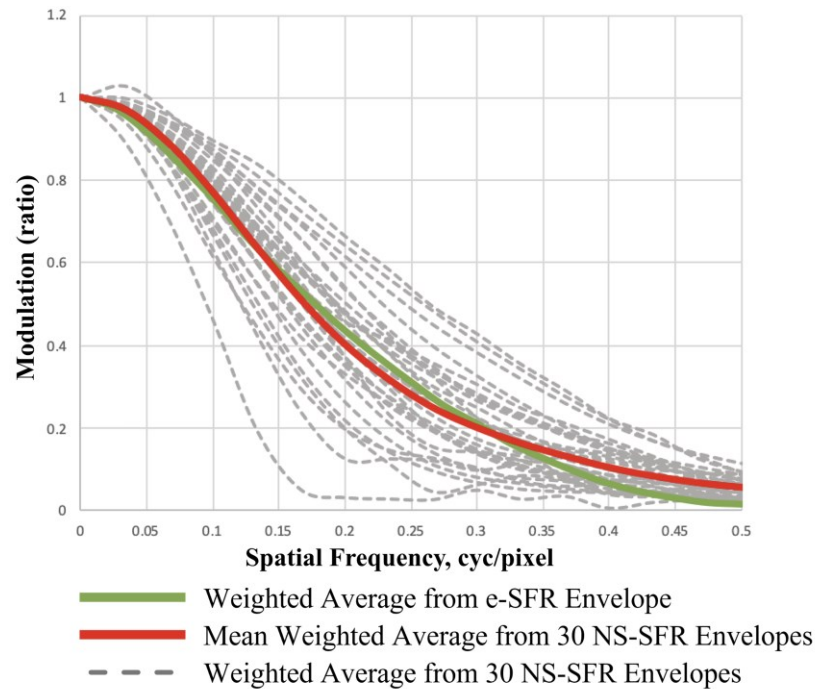


Figure 3.18 The mean of the weighted average NS-SFR envelopes of 30 natural scenes compared to the mean of a test chart e-SFR envelope [76].

3.7 Summary

This chapter has given an in-depth account of the proposed methodology designed to measure NS-SFRs from pictorial natural scenes, justifying each process used and highlighting the areas that would benefit from further development.

In summary, the edge locations in natural scene captures are first extracted using the Canny edge detector [123]. The resultant edge locations are used to crop the regions of interest (ROIs) of appropriate window sizes. The pixel stretching filter was developed to improve edge isolation and thus increase the yield of ROIs from a scene. Further benefits of this filter are decreasing the bias of image noise and illumination non-uniformity on the e-SFR. Each ROI goes through a process to validate that it contains a step-edge. Where appropriate, the ROIs are segmented into smaller ROIs to remove areas with non-uniformity streaking and split an edge of interest if the edges have a drastic direction change. The selected ROIs are then placed through the standard ISO12233 slanted-edge algorithm to measure the NS-SFRs. Burns' *sfrmat4* is utilised for this purpose [87].

An initial implementation of this framework was applied to 30 natural scenes, chosen for being well suited to this method. The results showed that the measured NS-SFRs strongly correlate with the ISO12233 e-SFRs. But, in contrast to laboratory-based measurements, the system and scene dependencies of the NS-SFR measurements are not controlled, resulting in NS-SFR envelopes that vary from scene to scene.

Assuming the system is linear, the system and scene dependencies that impact the NS-SFR envelopes are classified as:

NS-SFR system dependencies

- The PSF of the system hardware, including the sensor, optical system, and aperture (cf. § 1.1).
- The ROI edge location in the field of view (due to the natural variation of optical performance across the imaging circle) (cf. § 2.4.3).
- The optical depth of field controls how much of the scene is in focus (cf. § 2.9).
- The focus distance of a lens impacts system performance. Natural scene captures have a range of focus distances (cf. § 2.9).
- System image noise from the system hardware, introducing bias to the NS-SFRs (cf. § 2.10).

NS-SFR scene dependencies

- Unsuitable step-edge inputs, underestimating the e-SFR as they contain low modulation across the camera spatial frequency bandwidth (cf. § 2.9 & 3.5).
- High-frequency scene texture, introducing bias to the NS-SFRs (cf. § 3.5).
- Illumination changes around an edge of interest, forming non-uniformity or false sharpening (cf. § 3.3.2.3).

This separation between system and scene dependencies overlaps when measuring the NS-SFR from a non-linear system. Non-linear sharpening and denoising will depend on the local scene content (cf. § 2.4.2), which will locally affect the PSF. In addition, many modern systems use illuminance thresholds to set the strength of various processes, adding further scene dependency; for instance, lower light conditions will require higher noise reduction.

The following two chapters aim to estimate the system e-SFR performance from NS-SFRs.

Chapter 4 Sources of Variation in the Edge-Based Spatial Frequency Response

Sources of variation in the standard e-SFR must be first understood to estimate the system e-SFR from the NS-SFR data. Natural scene ROIs contain an assortment of edges with different parameters, introducing large variations in the data. Hence, these must be known and given consideration when developing the estimation methodology (cf. § 5.1).

This chapter details research into the ISO12233 e-SFR input parameters, expanding upon previous publications (cf. § 2.4.1) and applying application-specific variables, including pixel stretching (cf. § 3.3.2). The study uses simulated and captured edges to determine the variation introduced when changing single and multiple parameters. The aim is to provide a range of edge characteristics to obtain a stable system e-SFR estimation without excessively restricting NS-SFR data within which the ideal step-edges are rare.

4.1 Edge and ROI Parameter Variation

As reviewed (cf. § 2.4.1), the input edge and ROI parameters impact the slanted-edge method, causing variation in the e-SFR [5, 61, 63, 64]. The ISO12233 standard [5] minimises this variation by controlling the edges through a test chart captured in controlled laboratory environments. The established ranges/values:

- i) Edge angle of approximately 5° [5].
- ii) Edge Contrast between 0.55 and 0.65 Michelson Contrast [5].
- iii) ROI window width greater than 64 pixels [63, 69].
- iv) ROI window height between 80 to 500 pixels [63, 64].

When the slanted-edge algorithm is implemented using natural scenes, the parameters are no longer contained within these ranges, and suitable step-edge profiles are rare. If the established parameters, as mentioned above, were kept, few ROIs would be available for the e-SFR estimation. Furthermore, those selected edges cannot be guaranteed to be suitable for the SFR measure, i.e. contain modulation higher than the rendering capabilities of the system

(cf. § 2.1.3 & 2.9). The parameter values need to be re-evaluated to establish a selection range that isolates stable NS-SFRs for the system e-SFR estimation. The question now becomes, to what extent can the parameter ranges be set to provide good estimation while incorporating as many natural scene edges as possible.

4.2 Variation Study Methodology

This study builds upon previous publications that evaluated the slanted edge method [28, 61–64, 69] (cf. § 2.4.1), many of which follow similar methods. A series of ROIs are typically simulated or captured using a characterised camera system, and a single parameter is adjusted to evaluate the variation introduced. Image noise or processing may also be added to the test ROIs.

This study followed a similar approach; the objective was to benchmark the e-SFR variation with attributes specific to NS-SFRs, including edge and ROI parameters, listed below, pixel stretching edge isolation (cf. § 3.3.2) and image noise (cf. § 4.2.3). The Mean Absolute Error (MAE) was calculated between e-SFRs from an edge with the recommended parameters and an edge with varied characteristics to assess the variation (cf. § 4.2.4).

The tested parameter ranges are as follows:

- Ten edge angles were used, ranging between 5° to 40° , with 5° intervals. As 0° and 45° cannot produce a resampled ESF (cf. § 2.4.1), 2.5° and 42.5° were used to evaluate variation close to these ineffective angles.
- Thirteen edge contrast values were used, ranging between 0.07 and 0.92 Michelson Contrast, with intervals of 0.07.
- Seventeen ROI window heights were used, ranging from 20 to 180 pixels, with an interval of 10 pixels. An eighteenth ROI height was isolated at 128 pixels, a commonly used ROI height for benchmarking the e-SFR [22, 63, 134], providing the base height, i.e. a constant height for assessing the other parameters.
- Sixteen ROI window widths were used, ranging from 10 to 85 pixels, with an interval of 5 pixels. As with the ROI height, an additional ROI width was used as the base width, also set at 128 pixels. These base ROI height and width provide square ROIs to assess the steep edge angles. ROI widths between 85 and 128 pixels were not measured because narrow ROIs were of particular interest in isolating step-edges from natural scenes.

4.2.1 Simulated Step-Edges

A \tanh function, see Equation 4.1, was used in conjunction with the MATLAB function `meshgrid` [138] to simulate the step-edge 2D ROIs. The ROI window size, edge angle, contrast and edge gradient were controlled independently. These simulated step-edges were passed through `sfrmat4` code [87] to establish the resulting e-SFR.

The \tanh function simulates an ESF profile, $ESF_{sim}(x)$, across distance x , expressed by [60]:

$$ESF_{sim}(x) = \tanh(fx) \left(\frac{E_{max} - E_{min}}{2} \right) + \left(\frac{E_{max} + E_{min}}{2} \right) \quad (4.1)$$

where f controls the frequency content of the edge, i.e. the slope of the ESF profile, and E represents the maximum and minimum pixel values on either side of the edge.

This step-edge simulation technique was implemented to obtain 4,420 ROIs, each containing a different combination of the parameters of interest (cf. § 4.2). The mean pixel value was kept at 0.5 throughout the various contrast simulations. Not every combination of the edge parameters was required; this study focused on the following three:

- i) a base ROI window size of 128x128 pixels, with the entire angle and contrast value ranges (as demonstrated in Figure 4.1),
- ii) a fixed ROI height of 128 pixels with all parameter combinations of angle, contrast, and ROI width (as demonstrated in Figure 4.2),
- iii) a fixed ROI width of 128 pixels with all parameter combinations of angle, contrast, ROI height and noise values (as demonstrated in Figure 4.3).

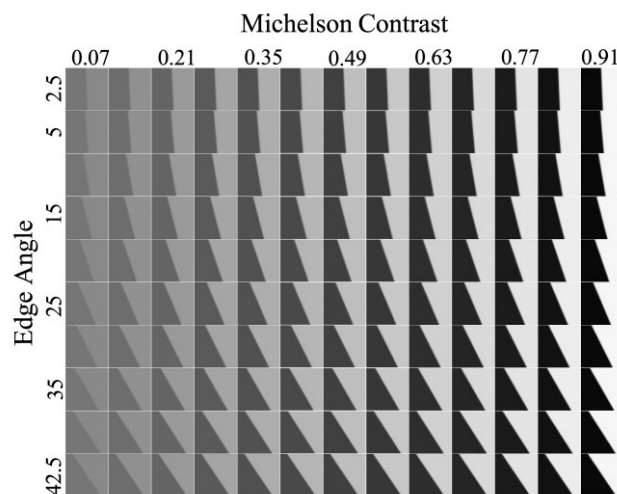


Figure 4.1 Examples of simulated square ROIs (128x128 pixels) ranging in angle and contrast. These examples are noiseless and without pixel stretching.

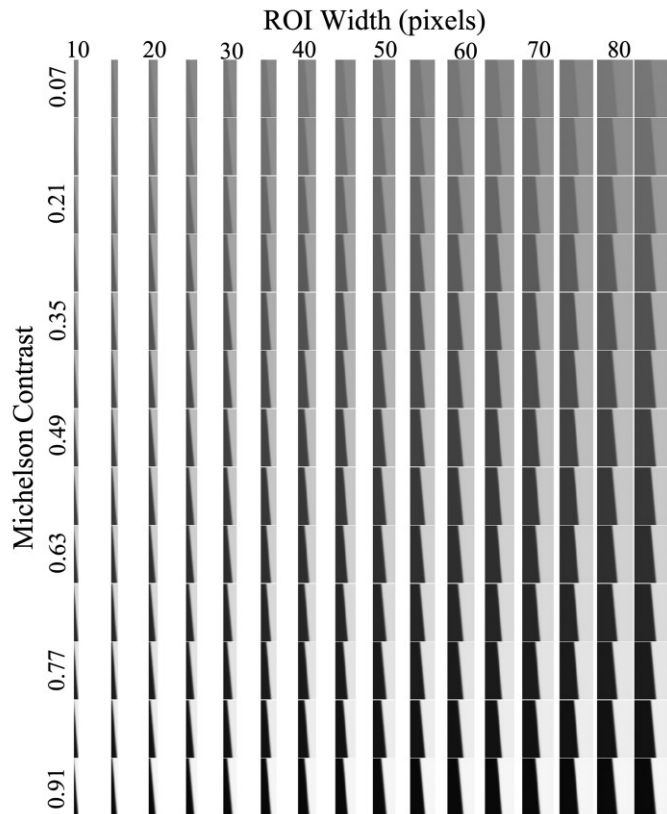


Figure 4.2 Examples of simulated ROIs with the height of 128 pixels ranging in contrast and ROI width. These examples are noiseless and without pixel stretching.

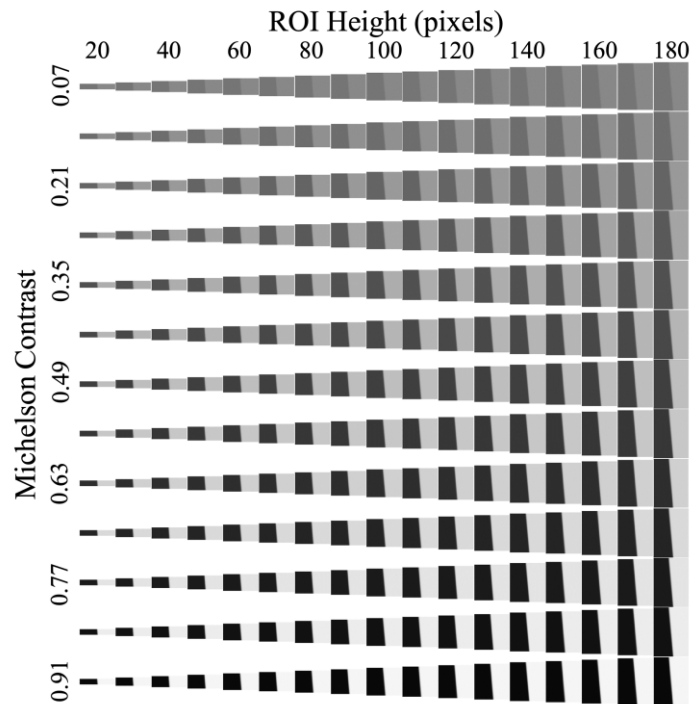


Figure 4.3 Examples of simulated ROIs with the width of 128 pixels ranging in contrast and ROI height. These examples are noiseless and without pixel stretching.

Six different SNRs were simulated on each of the ROIs, as described in depth in Section 4.2.3, which provided seven noise levels (including a noiseless iteration). The noise was applied three separate times to obtain an MAE for assessing variation (cf. § 4.2.4), which resulted in three simulated ROIs per parameter combination. This was followed by creating two versions of each ROI, one without and the other with pixel stretching (cf. § 3.3.2).

Using the standard slanted edge algorithm [5], the e-SFRs for each of these ROIs were obtained. In total, 167,960 simulated ROIs were created.

4.2.2 Step-Edge Test Chart Captures

The DSLR 1 system (cf. § 5.3.1) was used to obtain the captured step-edge test chart data. All images were captured in RAW, then converted to TIFF files, disabling sharpening and denoising in the demosaicing process. Images were also linearised by implementing the ISO14524 [2] as required by the slanted edge method [5] (cf. § 2.3.1). The resultant OECF can be found in Appendix A.

When capturing the test charts, the goal was to obtain the same ROI parameter combinations as the simulated data. This was achieved by using thirteen step-edge targets across the range of contrasts. They were photographed under the required controlled conditions dictated by the ISO12233 (cf. § 2.3), the lighting on the test chart stayed within a 2% illumination variation [5], and the test charts and camera sensor was at an appropriate distance apart [49]. Each test chart was rotated at set markings to obtain the desired edge angle range, maintaining the pivot point at the centre of the camera system field of view. Figure 4.4 provides an example of three of the image captures.

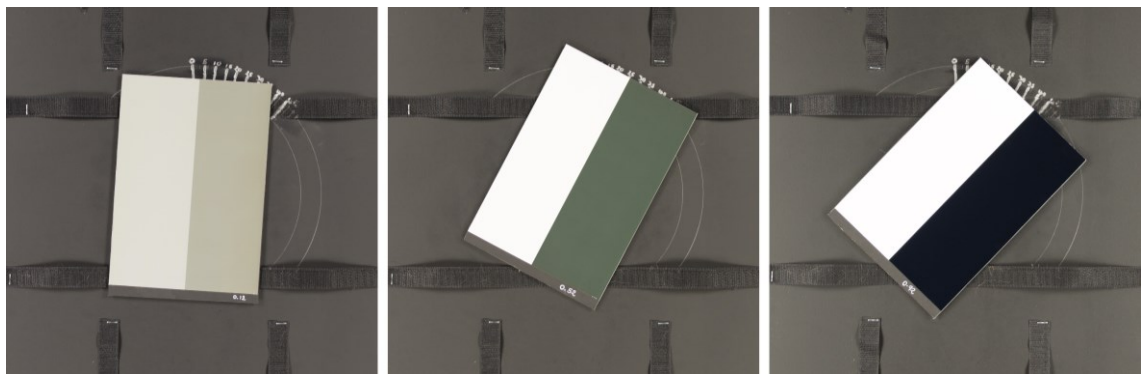


Figure 4.4 Examples of a captured step-edge test chart, at Michelson contrasts of 0.12, 0.52 and 0.92, positioned at 2.5°, 25° and 42.5°.

The system ISO gain setting was adjusted to obtain the range of image noise levels; the shutter speed was adjusted to maintain a constant exposure when changing the ISO. Each test chart orientation was captured six times, at ISO100, ISO200, ISO400, ISO800, ISO1600 and ISO3200, resulting in 780 images.

These images were then cropped to the appropriate ROI window sizes, obtaining the same three parameter combinations as the simulated data. Although close, the parameter range intervals were not as precise as intended; for example, the edge angle rotation was subject to human error. That said, the angles were all within 4 degrees of the desired angle. In addition, the edge contrast range did not have even intervals between 0.07 and 0.98 Michelson Contrast because the printed charts were not correctly scaled to the linearisation of the images. The variation assessment has addressed this (cf. § 4.2.4).

Again, three ROIs were required for each parameter combination to obtain an MAE, achieved by cropping a ROI from the centre, one above and one below, keeping the edge central to the ROI window. There were no noiseless images; thus, there were fewer captured ROIs than the simulated data. In total, 159,120 test chart captured ROIs were cropped.

4.2.3 Applying Image Noise

To achieve the same SNR values as the captured data, first, the ISO15739 [4] (cf. § 2.10) was implemented to obtain the SNR of each of the six DSLR 1 (cf. § 5.3.1) ISO gain settings, measured from 13% of the saturated exposure as per the ISO15739 [4]. The results were SNR values 97, 76, 51, 37, 27 and 18, corresponding to ISO settings from 100 to 3200.

A combination of Poisson and Gaussian distributions was used to model shot and read noise, respectively, as described in the literature [45, 48, 121] (cf. § 2.10) for application to simulated images. In practice, the simulated ROI was first converted from pixel values to electrons (ROI_e), achieved by normalising the ROI pixel values, Equation 4.2, and multiplying the normalised ROI by the simulated well capacity (i.e. the number of electrons equal the saturated white level), Equation 4.3.

$$ROI_{norm}(x, y) = \frac{ROI(x, y)}{b} \quad (4.2)$$

$$ROI_{e-}(x, y) = ROI_{norm}(x, y) \cdot w \quad (4.3)$$

where b is the ROI bit-depth and w is the well capacity. The well capacity was determined by dividing the *full well capacity* (FWC), FWC , by the desired gain value, g , Equation 4.4.

$$w = \frac{FWC}{g} \quad (4.4)$$

Figure 4.5 illustrates this process [139]. A higher gain effectively means a smaller FWC.

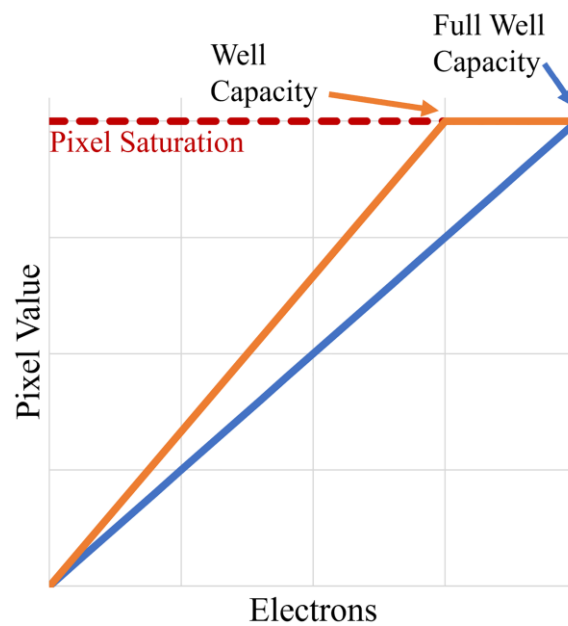


Figure 4.5 The relationship between the number of electrons and the pixel value. The trend is linear for a gain value of 1 (no gain), reaching the FWC. When the gain is applied, the well capacity decreases, requiring fewer electrons to reach pixel saturation. Adapted from [139].

The amount of shot noise applied is equal to the square root of the number of quanta in the pixel [45] (cf. § 2.10). If one quanta, q , is assumed to be equal to one electron, this yields shot noise of \sqrt{q} . Using the Poisson distribution shot noise was then emulated across ROI_{e-} , as expressed as [140]:

$$ROI_{e-shot}(x, y) = \frac{1}{\lambda} Poisson(\lambda ROI_{e-}(x, y)) \quad (4.5)$$

were

$$\lambda = \sqrt{q} = \sqrt{ROI_{e-}(x, y)} \quad (4.6)$$

Read noise was generated using a Gaussian distribution, with a standard distribution, σ , of 3 electrons and a mean value, μ , of zero. At each pixel, the noise term, $ROI_{READ}(x, y)$ (a pixel array with the mean value μ), was characterised by a probability of the Gaussian distribution, p , as expressed in Equation 4.7, [141].

$$p(ROI_{read}(x, y)) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-ROI_{read}(x, y)^2}{2\sigma^2}\right)} \quad (4.7)$$

These noise models were achieved by implementing the MATLAB *imnoise* function [142]. The sum of the resulting noise arrays, the shot noise with the signal and the read noise, provided the ROI_e - with the total noise.

$$ROI_{e-total}(x, y) = ROI_{e-shot}(x, y) + ROI_{read}(x, y) \quad (4.8)$$

The gain value was then multiplied with noisy ROI_e - to obtain the desired signal level. The noise is also amplified in this process.

$$ROI_{e-gain}(x, y) = ROI_{e-total}(x, y) \cdot g \quad (4.9)$$

The resulting amplified ROI_e - was then converted back to pixel values.

$$ROI_{noise}(x, y) = ROI_{e-gain}(x, y) \cdot b \quad (4.10)$$

Changing the gain value in this model results in different SNRs. Increasing the gain results in a decreased signal, which in turn decreases the SNR. The SNR was measured by passing a 13% grey patch through the same noise model as the simulated ROIs. Rather than combine the shot and read noise to the signal in the model, the signal and total noise were kept separate, allowing a straightforward measurement of the signal (13% of the well capacity) divided by the total noise variance, Equation 2.43.

To obtain the six SNRs, equivalent to the DSLR 1 ISO gain settings, an FWC of 3000 electrons were used with the gain values of 1.05, 1.76, 2.90, 4.54, 6.85 and 10.63. Figure 4.6 provide examples of this simulation.

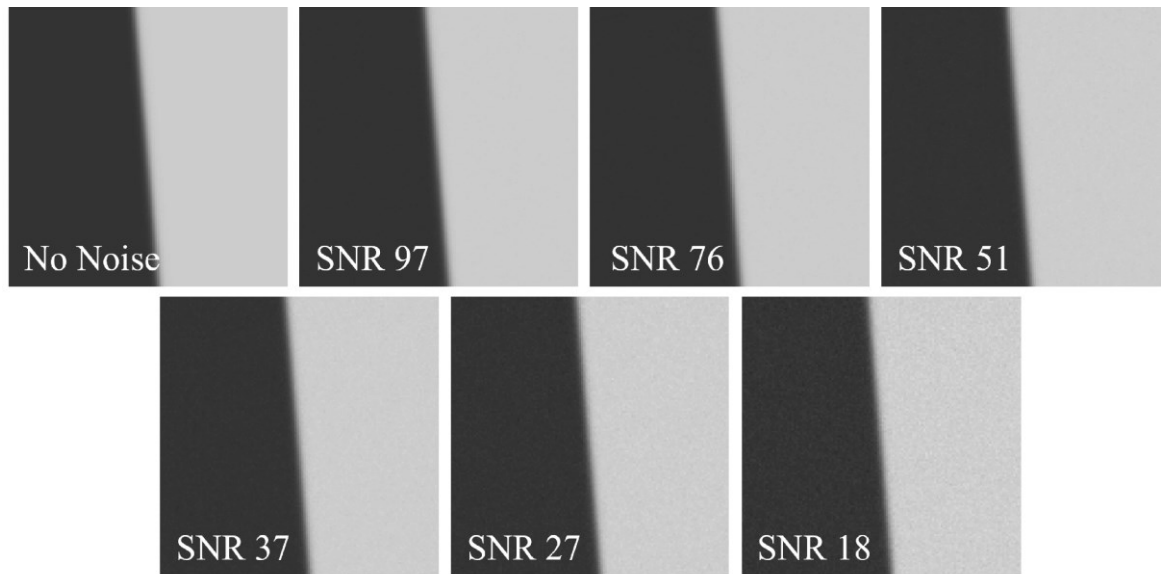


Figure 4.6 A simulated 5° step-edge in a 128x128 pixel ROI window, with the six levels of image noise emulation used in this study.

4.2.4 Variation Assessment

As mentioned, the MAE was used to evaluate the error between a reference e-SFR measured from a ROI containing the established recommended parameters and each e-SFR derived from the parameter combinations. The reference e-SFR parameters were a 5° angle, a Michelson contrast of 0.60 and a window size of 64x128 pixels (cf. § 2.4.1).

The MAE measurements were obtained by first using *Piecewise Cubic Hermite Interpolating Polynomial* (PCHIP) [143, 144] to resample the e-SFRs to 51 samples at spatial frequencies ranging from 0 to 0.5 cyc/pixel, providing equivalent sample points to calculate the absolute errors. Every parameter combination was simulated/captured three times; the mean of the three MAEs was computed to provide the average variation.

The resulting MAEs were assigned to a six-dimensional (6D) coordinate system. This coordinate system was sized 10x13x17x16x51x6, with dimensions angle, contrast, ROI height, ROI width, spatial frequency and SNR. Note, the simulated data has a seventh value for the SNR coordinate, 10x13x17x16x51x7, that contained the noiseless ROIs.

This process was repeated four times, producing coordinate systems describing the variation from the simulated and captured test chart data, with and without pixel stretching being applied.

4.3 Multi-Edge Parameter Variation Results

This section examines the MAEs from the 6D coordinate system to determine appropriate ranges for the system e-SFR estimation. First, the simulated data is discussed (cf. § 4.3.1), followed by the captured data (cf. § 4.3.2).

4.3.1 Simulated e-SFR Results

Figure 4.7 illustrates the MAE from the four individual parameters. The base settings are a 5° edge angle, a Michelson contrast of 0.60 and a ROI window size of 64x128 pixels; each parameter is exercised over its full range (cf. § 4.2) whilst keeping the others constant.

The mean MAE is calculated for the high frequencies (0.37 to 0.50 cyc/pixel) to determine the effects of error from image noise on parameter variation. The MAE is calculated twice, for the data with and without pixel stretching.

This data highlights how pixel stretching reduces the effects of noise on the e-SFRs. Without the edge isolation, noise effects are apparent and greatest for low angles, low contrasts, small ROI heights, and wide ROI widths. Variation is significantly reduced with pixel stretching, which provides results closer to the expected noiseless behaviour (cf. § 2.4.1 & 3.3.2.1).

Figures 4.8 and 4.9 show matrices containing plots of the simulated data in colour mapped scatter plots, depicting multi-parameter variation without pixel stretching. Figure 4.8 plots the results without noise, illustrating only the parameter variation behaviour. Figure 4.9 shows the variation and impact of image noise, using data with SNR 18. The colours depict the mean MAE calculated between the frequencies 0.0 to 0.5 cyc/pixel, which provides the average variation behaviour of the entire e-SFR. Finding in these matrices further support the literature (cf. § 2.4.1), supporting the expected behaviour:

- an increase of error with edge angle,
- low contrast edges prone to error from image noise,
- a slight decrease in error with the increase of ROI height,
- narrow ROIs impede the ESF introducing error, whilst the wide ROIs show an increase of error from image noise.

It is also observed that parameters prone to image noise can compile to a further increase the variation, the prominent parameters being the combination of low contrast and short ROI heights.

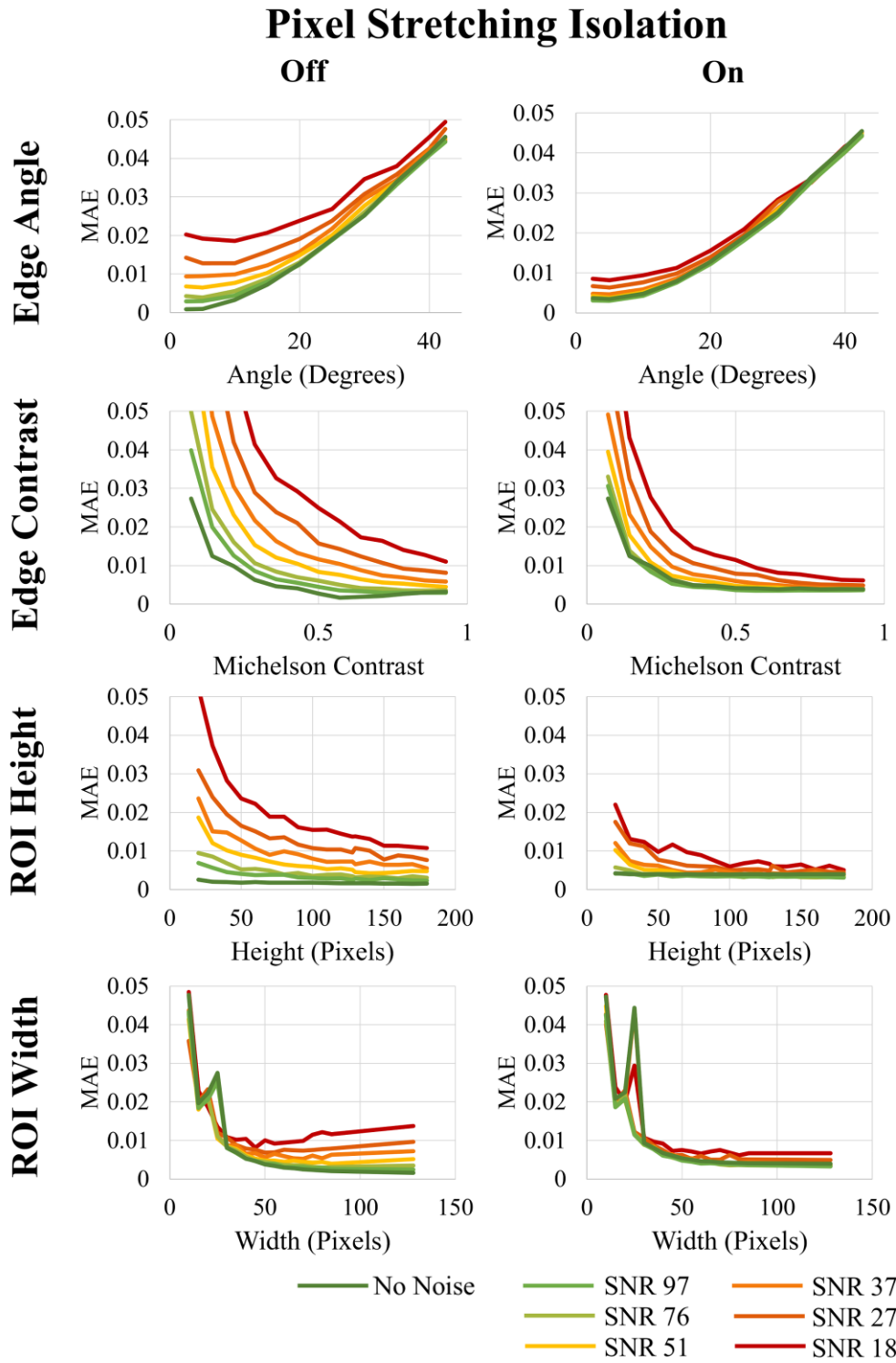


Figure 4.7 Comparison of the simulated MAE introduced by the SFR parameter ranges, edge angle, edge contrast, ROI height and ROI width, between spatial frequencies 0.37 to 0.50 cyc/pixel, with and without applying pixel stretching at seven noise levels.

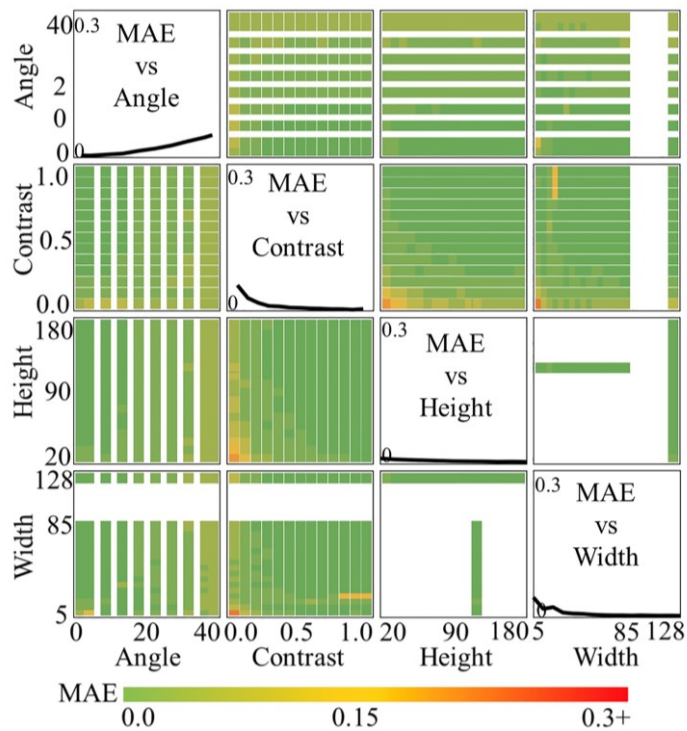


Figure 4.8 Matrix of simulated noiseless variation introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by the individual parameter ranges.

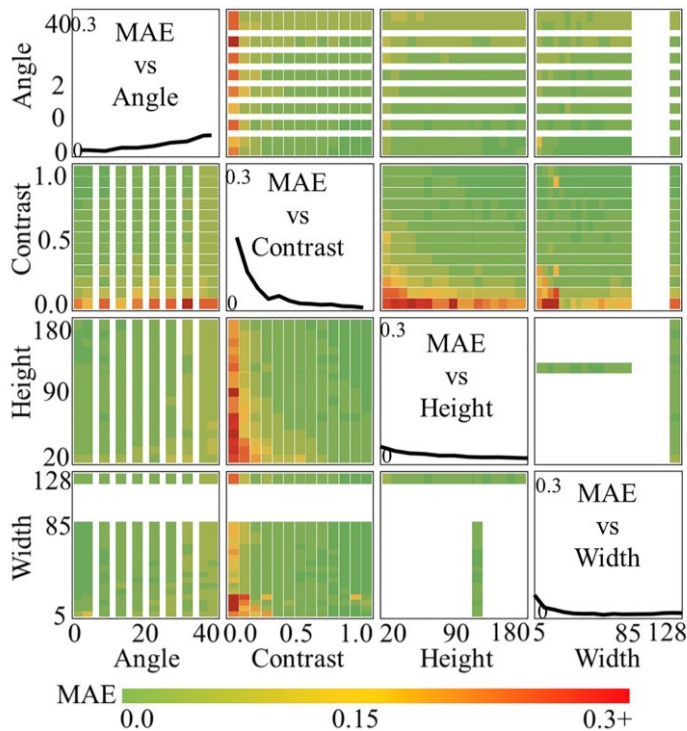


Figure 4.9 Matrix of simulated SNR 18 variation introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by the individual parameter ranges [145].

4.3.2 Captured Test Chart Results

Figures 4.10 and 4.11 illustrate the captured test chart results, plotted in the same manner as the simulated data. Many observations are shared between the simulated and captured test chart data, including the positive image noise reduction effects of pixel stretching, as seen in the low contrast edges, the ROI height, and the wide ROI widths. The ROI window size is shown to introduce negligible MAE, especially when pixel stretching is applied, providing the grounds to use small ROIs in the e-SFR estimation without introducing variation.

Observations from the DSLR 1 system indicate inconsistencies compared to the simulated data. Although sharpening and denoising were switched off in the demosaicing process, MAE variation in the output indicates non-linear behaviour in the system pipeline. It is apparent in the contrast variation range in Figure 4.10; there is an increase in MAE with high contrast edges, unaffected by pixel stretching (i.e. it is not a noise related error). No sharpening lobe is observed when directly evaluating the e-SFRs, whilst a positive error appears only in the high frequencies. The edge angle also behaves unexpectedly. Rather than the anticipated MAE trend of an increase in MAE with edge angle, there is a peak around 22.5° ; again, this error is only in the high frequencies. It is speculated that these two effects are caused by the demosaicing algorithm. The anisotropic nature of the sensor array may result in the demosaicing interpolation having non-linear spatial effects. This may lead to better interpolation of near-vertical edges and edges at 45° . At 22.5° , it is likely to become less efficient, introducing error to the e-SFR. Also, it could be more effective with higher contrast edges, altering the output e-SFR.

Jenkin provided a mathematical model that describes how the recorded modulation changes with respect to the skewness of a target [20, pp. 97–104]. Figure 4.12 provides an example of the optimal recorded modulation for vertical and 45° edge orientations.

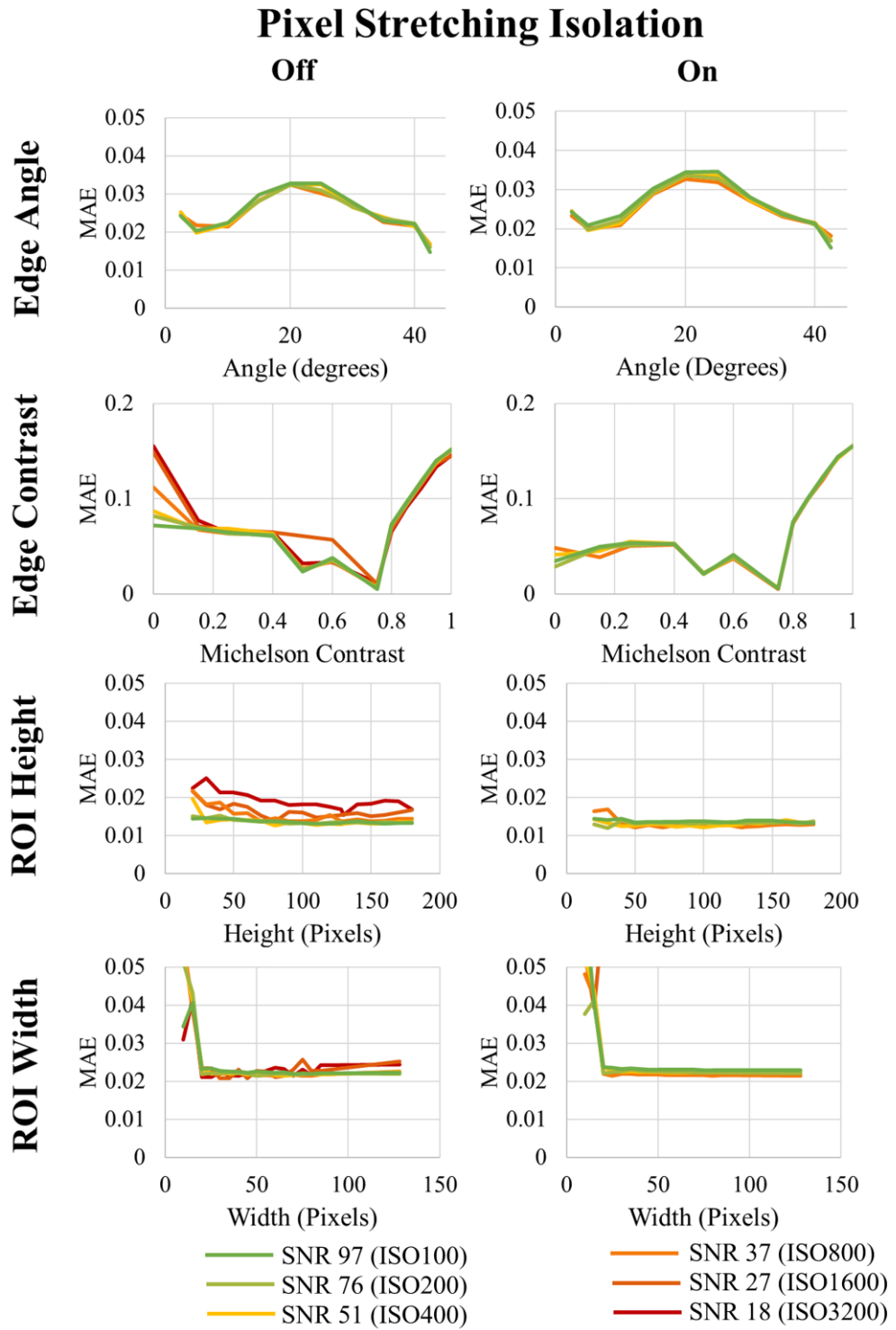


Figure 4.10 Comparison of the DSLR 1 captured test chart MAE introduced by the SFR parameter ranges, edge angle, edge contrast, ROI height and ROI width, between spatial frequencies 0.37 to 0.50 cyc/pixel, with and without applying pixel stretching at six SNRs.

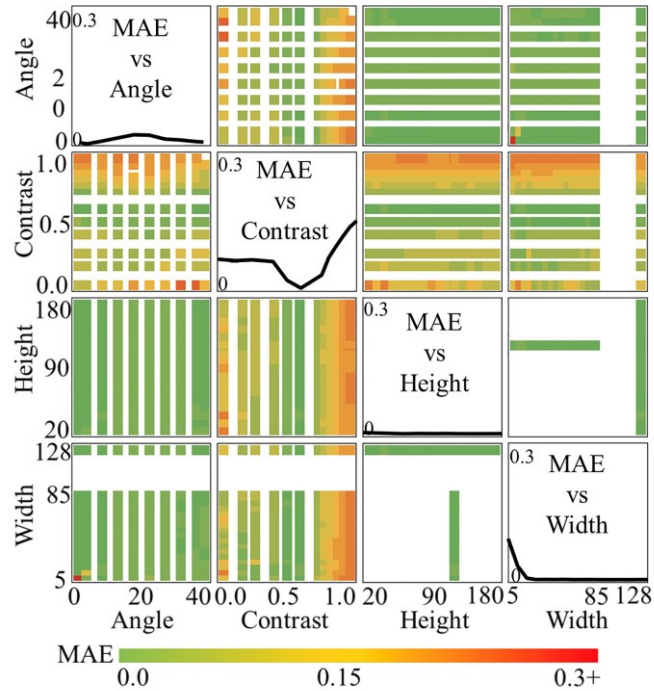


Figure 4.11 Matrix of variance introduced to the slanted-edge e-SFR measurement when changing edge angle, edge contrast, ROI height and width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots the MAE introduced by each SFR parameter. The data is from test chart ROIs, captured using the DSLR1 at a gain setting of ISO100 (SNR 97).

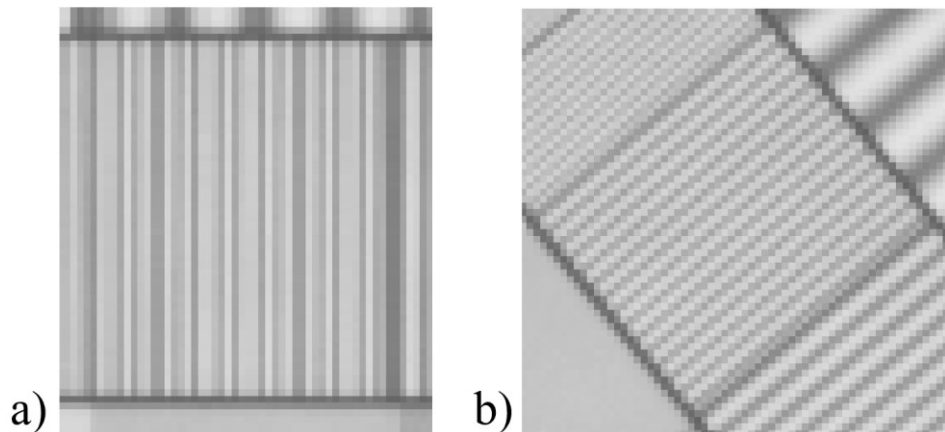


Figure 4.12 The optimal edge angles (a) 0° and b) 45°) for recording the modulation with a sensor array, demonstrated with a sine-wave test chart with the frequency of 0.27 cyc/mm. Adopted from [20].

4.4 Parameter Selection for System e-SFR Estimation

Table 4.1 lists NS-SFR parameter ranges established for estimating system e-SFR (cf. § 5.1). There aim to maximise the number of edges utilised from a natural scene without introducing excessive NS-SFR variation. These ranges are compared in the table to 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 by reducing the image noise, it does not affect the edge itself, thus will not change any non-linear ISP applied. For this reason, the contrast range was kept from 0.55 to 0.65 Michelson Contrast, as per ISO12233, to reduce the effects of non-linear sharpening.

Variation from the ROI window size is noise dependent and is reduced by implementing pixel stretching; therefore, a broad range of ROI sizes was chosen to be implemented in the estimation. When cropping edges from natural scenes, tall ROIs are more likely to contain unwanted artifacts, such as depth of field changes across the edge, intersecting edges and illumination non-uniformity. Therefore a maximum ROI height was set at 128 pixels, a typical small ROI height for the slanted edge algorithm [22, 63, 134]. The NS-SFR framework ROI crop sizes (cf. § 3.3) were set from the parameter ranges established here, so that the NS-SFR are isolated at appropriate window sizes, and no restrictions are required in the system e-SFR estimation (cf. § 5.1).

Parameter	ISO 12233 e-SFR	NS-SFR
ROI Size	>64 x 80-500 pixels	>20 x 20-128 pixels
Edge Angle	<45° (5° Recommended)	2.5° – 35°
Edge Contrast	0.55 – 0.65 Michelson Contrast	0.55 – 0.65 Michelson Contrast

Table 4.1 ROI and edge parameter ranges for the e-SFR ISO 12233 method against those used to estimate the system-SFR from the NS-SFR data.

4.5 Summary

This chapter has presented a study on the variation introduced to the e-SFR by changes in edge angle, contrast, ROI height and width across a relevant SNR range. The research builds upon previously established publications [28, 61–64, 69] (cf. § 2.4.1) that have benchmarked the e-SFR in terms of their input parameters. The objective was to determine parameter ranges to estimate an accurate system e-SFR from NS-SFR data. This should be done without restricting the yield of natural scene step-edges, from which suitable step-edge inputs are rare.

The method used both simulated and test chart captures to obtain edges with a range of parameters. Three e-SFRs were calculated per unique parameter combination to compute an MAE. The e-SFRs were measured with and without pixel stretching (cf. § 3.3.2), resulting in 167,960 and 159,120 e-SFRs for the simulated and captured edges, respectively. The MAEs were placed into a 6D coordinate system, where each coordinate represented a parameter combination of angle, contrast, ROI height, width, spatial frequency and SNR.

The results were presented in two formats. The first plotted the individual parameter variation per SNR, with and without pixel stretching. The second plotted the MAE in a colour-coded matrix that illustrates the variation relationship between multiple parameters. Due to the traits of pixel stretching, the error from image noise was significantly reduced. With this noise reduction, broader parameter ranges can be used without noise becoming a decisive factor.

Although sharpening and denoising in the DSLR camera system TIFF pipeline were disabled, the MAE data from the captured edges indicated non-linear behaviour. This is assumed to be caused by the anisotropic effects of the sensor array and the interpolation algorithms used in the demosaicing process, but is to be confirmed with further work.

The research provided grounds to establish the parameter ranges to estimate the system e-SFR from NS-SFRs, as follows:

- The angle range was set to be 2.5° to 35° . Although this range introduced some variation to e-SFRs, it does not restrict the step-edge yield from captured natural scenes, from which a broad range of edge angles can be found.

- The contrast was kept at a mid-range, between 0.55 to 0.65 Michelson Contrast. Low contrast edges are prone to noise bias, and high contrast edges are susceptible to non-linear sharpening. That said, if required in certain applications, pixel stretching allows the edge contrast range to be increased to lower contrast edges without introducing bias from image noise.
- The ROI window size was set at widths greater than 20 pixels and a height between 20 to 128 pixels.

The angle and contrast are set as thresholds in the estimation methodology (cf. § 5.1), whilst the ROI window size was programmed into the NS-SFR framework (cf. § 3.3).

Chapter 5 Estimation of the System Spatial Frequency Response

This chapter explores the estimation of the ISO12233 system e-SFR (cf. § 2.3) from NS-SFR data gathered from the proposed framework (cf. § 3.1). Deriving the system e-SFR is based on binning the NS-SFR data in a multi-parameter coordinate system. The system e-SFR is estimated by averaging NS-SFRs within the predetermined parameter ranges (cf. § 4.4) per radial distance segment. The methodology is detailed, providing the rationale behind each step. The established coordinate system is also used to confirm that the edge parameter ranges offer the most stable data for the estimation, using the NS-SFRs rather than simulated and test chart captures.

The system e-SFR estimation method is implemented on diverse natural scene image datasets from three characterised camera systems, two (near-) linear and one non-linear. Quantitative analysis is carried out on the system e-SFR estimates to validate the accuracy and precision of the method.

5.1 System e-SFR Estimation Methodology

An overview of the methodology to estimate the system e-SFR from NS-SFR data is first outlined, briefly summarising the approach. Subsequent sections then provide further depth to the main elements of the methodology and its development.

5.1.1 Overview of the Methodology

The principle behind the system e-SFR estimation is to regulate the NS-SFRs. The data is selected and grouped to minimise variation introduced from both the scene and system (cf. § 2.9 & 3.6), which increases the probability of providing an accurate e-SFR estimate. This data is then averaged to determine the system e-SFR. The estimate is achieved by the following steps:

1. Data organisation (cf. § 5.1.2) – The NS-SFRs are grouped according to their edge orientation (horizontal or vertical) and their position in the field of view. In addition

to horizontal or vertical edge orientations, the data is converted into sagittal and tangential orientation groups.

2. Selection of *suitable* step-edge (cf. § 5.1.3), i.e. step-edges that contain modulation higher than the rendering capabilities of the system (cf. § 2.1.3 & 2.9) – The narrowest LSFs are selected, corresponding to the NS-SFRs that most likely deliver the highest performance, and thus the system e-SFR.
3. Data binning and system e-SFR estimation (cf. § 5.1.4) – The selected NS-SFRs are assigned to a 4-dimensional (4D) coordinate system, binning spatial frequency, edge angle, contrast and radial distance. The NS-SFR data from edges with the predefined angle and contrast (cf. § 4.4) are averaged per radial distance, deriving the e-SFR across the frame. A weighted average of these localised e-SFR estimations gives a global e-SFR system estimate.

This methodology is summarised in the Figure 5.1 flowchart. The 4D coordinate system, marked with C, can be used for NS-SFR analysis, for example, verifying edge parameter stability (cf. § 5.2).

5.1.2 Data Organisation

The NS-SFR data (cf. § 3.1) is first restructured, compiling the data from each image in the dataset into a cell array, segmented by radial distance frame partitions. This location segmentation reduces the optical e-SFR variation across the imaging circle (cf. § 2.4.3), providing a localised estimation of system e-SFR per radial distance. Six radial distances are used, 1/6 refers to the central image area, and 6/6 refers to the corners of the field of view. The number of radial distance segments can be changed depending on the application of the e-SFR. Horizontal and vertical edge orientations are kept separate due to the non-stationary nature of a sensor array (cf. § 1.1), i.e. performance is orientation dependant.

The horizontal and vertical edges are converted to provide sagittal and tangential orientations (cf. § 2.6.2), utilising the knowledge of the edge angles and locations in the field of view. The frame is divided into 16 diagonal subdivisions passing the frame centre. These subdivisions are given two edge angle ranges, determining whether an edge orientation is appropriate for sagittal or tangential. If true, the NS-SFR data is copied into a sagittal/tangential data array, again segmented by radial distance. The angle ranges used per subdivision are shown in Figure 5.2.

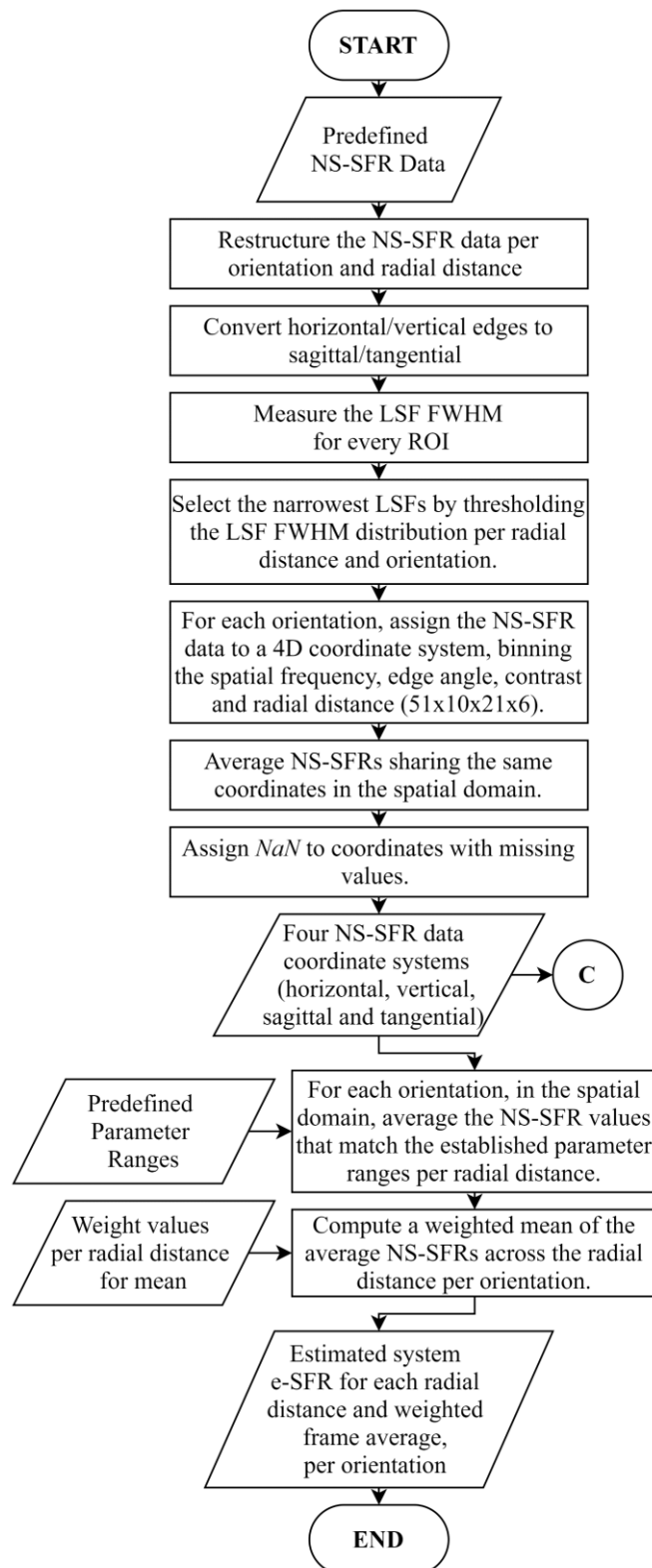


Figure 5.1 Flowchart depicting the proposed method to estimate the system e-SFR from natural scene derived SFR data.

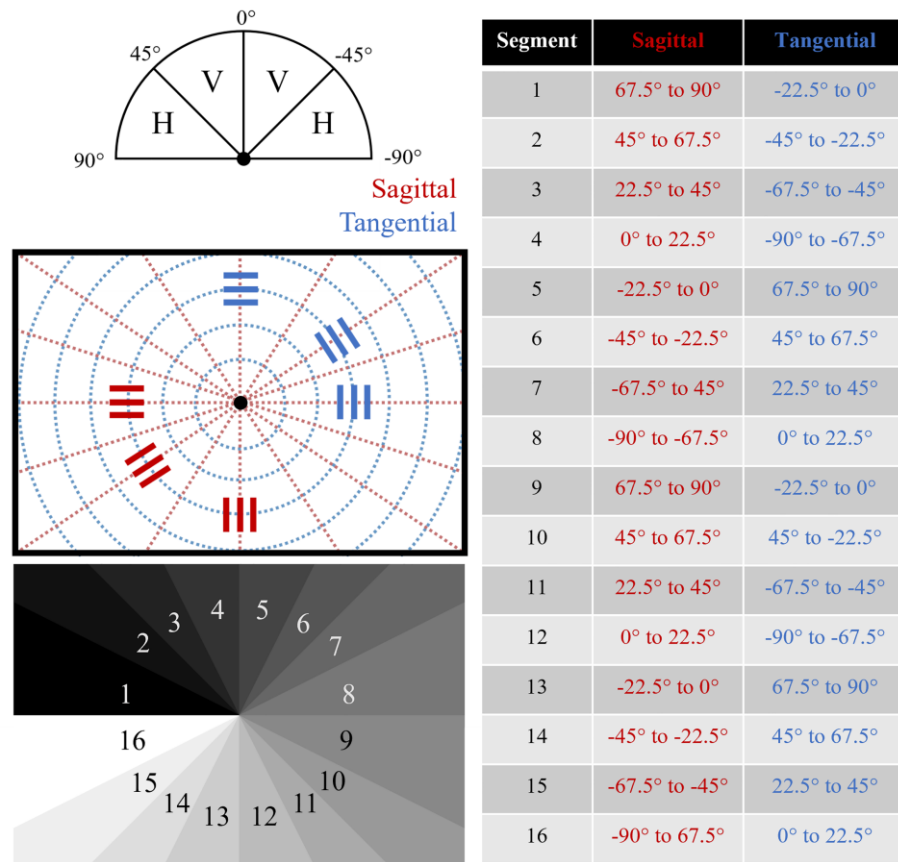


Figure 5.2 The horizontal and vertical angles were converted for each of the 16 diagonal frame subdivisions used to obtain the sagittal and tangential edges.

5.1.3 Selection of Suitable Step-Edges

To estimate system performance, the e-SFR must be measured from suitable step-edge inputs (cf. § 2.1.3 & 2.9). Therefore, a threshold is set to obtain the suitable step-edges from the NS-SFRs, i.e. selecting edge inputs that maintain higher modulation across the spatial frequencies than the camera system's rendering capabilities. This threshold is achieved by measuring the NS-SFR LSF FWHM distribution per radial distance and orientation, then selecting the narrowest LSFs. A threshold is applied to the LSF FWHM distribution to return the narrowest LSFs. The top 10th, 20th and 50th percentiles of distribution are compared for this purpose.

The 10th percentile threshold provides only the narrowest LSFs without restricting excessively the amount of edge data. The 20th percentile threshold expands the NS-SFR yield but still returns narrow edges. Finally, the 50th percentile threshold splits the isolated edges in half, using more NS-SFR data and a larger range of LSF widths.

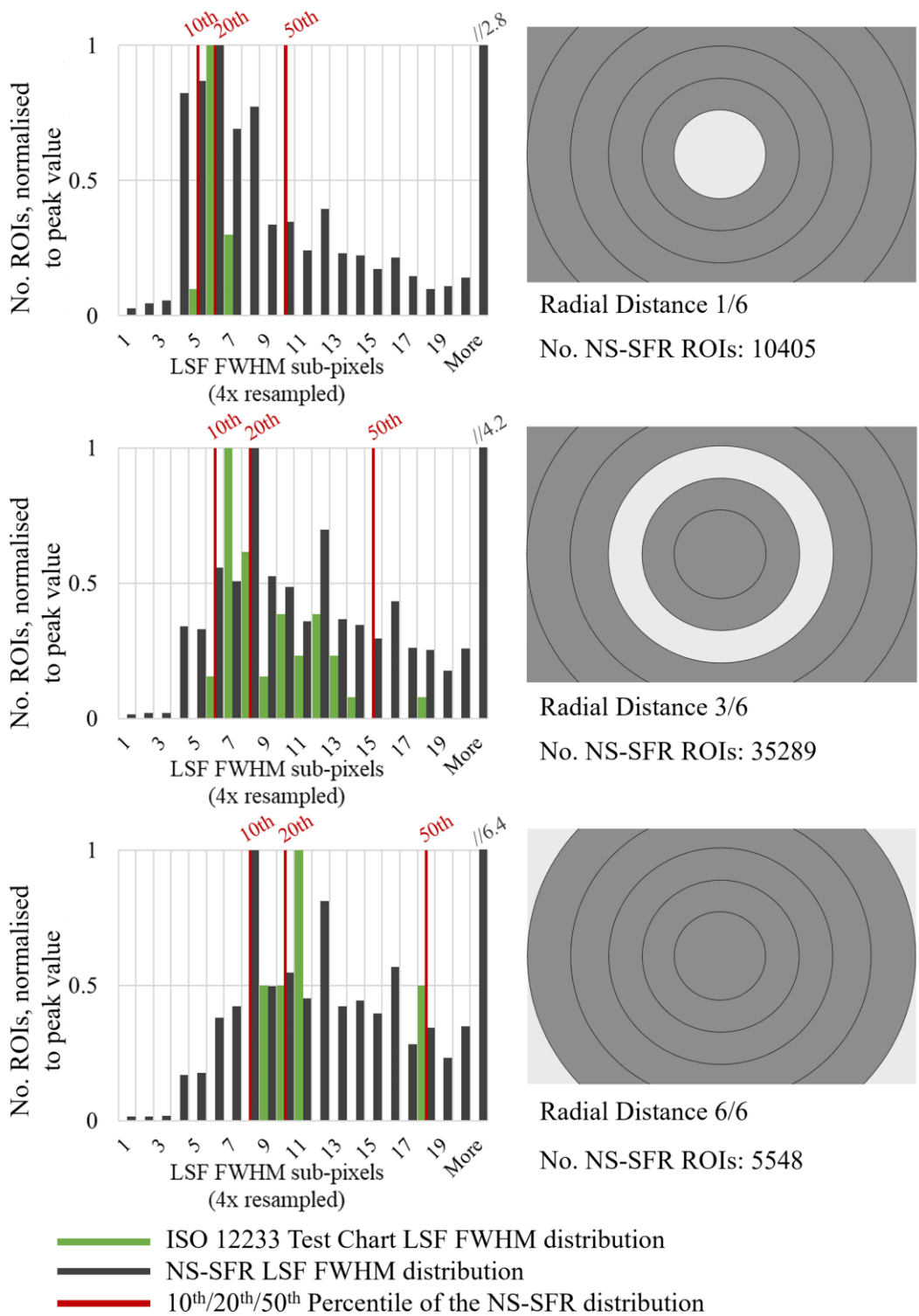


Figure 5.3 These histograms illustrate LSF FWHM from the test chart and NS-SFR step-edges for data from the DSLR 1 TIFF image dataset. The histograms are normalised using their peak value. This example illustrates distances 1, 2 and 6 out of 6 radial segments. The red bars indicate the 10th, 20th and 50th percentile for the NS-SFR LSF FWHM distribution [124].

Figure 5.3 depicts the DSLR 1 (cf. § 5.3.1) vertical NS-SFR LSF FWHM distribution for radial distances 1/6, 3/6 and 6/6. By comparing the position of the 10th percentile to the distribution from a test chart, it is seen that it corresponds to the narrowest ISO12233 test chart LSF output. Although expanding the threshold to 20th or 50th percentiles incorporate more of the ISO12233 e-SFR LSF FWHM distribution, this was shown to increase the probability of introducing unsuitable step-edge inputs (cf. § 5.3.3). Therefore, this input uncertainty must be considered when evaluating the system e-SFR estimation from each of these thresholds.

5.1.4 Data Binning and System e-SFR Estimation

In a similar way to the e-SFR variation study (cf. § 4.2), a coordinate system is used to organise NS-SFR data, this time to derive the system e-SFR. Previously established edge and ROI parameter ranges have shown that only NS-SFR angle and contrast should be restricted (cf. § 4.4). Thus, selected NS-SFR data is arranged into a 4D coordinate system; each coordinate refers to a unique combination of the following parameters: spatial frequency, edge angle, edge contrast and radial distance, with each coordinate having an associated NS-SFR value. The edge selection in the NS-SFR framework controls the ROI window size, allowing for the appropriate ROI sizes (cf. § 3.3).

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 NS-SFRs have the same coordinates, their LSFs are averaged in the spatial domain to reduce the high-frequency bias caused by image noise [39]. For averaging a series of LSFs, first, they are registered, aligning their peak values, then the area under each LSF is normalised, and the mean is calculated at each sample point. The average LSF is finally used to recalculate the NS-SFR for the given set of coordinates.

PCHIP interpolation [143, 144] is used to resample the spatial frequencies into 51 samples ranging from 0 to 0.5 cyc/pixel. 51 frequency values sample the NS-SFRs at selected summary measures (cf. § 2.6.1), such as 25%, 50% and 75% of the Nyquist frequency, as

well as providing high-resolution sample points for analysis. The number of samples can be adjusted for purpose.

These sets of NS-SFR values, binned at the given parameter coordinates and specified frequencies, are computed per radial distance, resulting in the 4D binning coordinate system, sized 51x10x21x6 (spatial frequency, edge angle, edge contrast and radial distance). Any coordinate without assigned data is set as *Not a Number* (NaN). Four coordinate systems were created, one for each orientation: vertical and horizontal, sagittal and tangential.

The binned NS-SFRs are then thresholded using the established parameter ranges in Table 4.1. This selected data is then averaged in the spatial domain per radial distance, resulting in six local system e-SFR estimates across the frame. A weighted average of these six estimates is calculated to obtain a global system e-SFR estimate. Again, the average is done in the spatial domain. The chosen weightings are 1.00 for the centre, 0.75 for the partway regions and 0.50 for the frame's corners, corresponding to default weights in Imatest software employed for the SFR analysis [74]. They 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 centre, emphasising the poorer SFRs [75] (cf. § 2.4.3).

As seen in Figure 5.3, 3/6 (partway) region of NS-SFR data contained three times the number of ROIs than the frame centre and more than six times compared to the frame corners. The weighted mean is calculated from the system e-SFR estimates of the radial distances rather than individual NS-SFRs, removing bias caused by radial segments areas of high-density edges.

5.2 Validation of the Established Parameter Ranges

The 4D coordinate system used to estimate the system e-SFR (the output at C in Figure 5.1) was utilised to validate the established parameter ranges (cf. § 4.4). A 5th dimension was added to the coordinate system, the ROI height, to confirm that the ROI window size introduces a negligible variation on the NS-SFR data. The ROI height ranged from 30 to 130 pixels, at intervals of 10, providing 11 values.

The resulting 5-dimensional (5D) coordinate system was sized 51x10x21x11x6 (spatial frequency, edge angle, edge contrast, ROI height and radial distance), with an NS-SFR value

per coordinate. Three 5D coordinate systems were calculated per dataset, setting the LSF FWHM distribution threshold at 10th, 20th and 50th percentiles.

Two datasets of NS-SFR were implemented, captured from the characterised DSLR 1 and DSLR 2 systems, as detailed later in this chapter (cf. § 5.3.1). DSLR 1 dataset had two iterations, the first from the RAW sensor images and the second from demosaiced TIFF images with sharpening and denoising turned off. The DSLR 2 data represented only the demosaiced TIFF pipeline.

The MAE was calculated between the systems' ISO12233 e-SFR [5] (cf. § 5.3.1) and the NS-SFRs; this was done per radial distance. As previously (cf. § 4.2.4 & 4.3.1), the spatial frequencies were binned to only use a specified range in the MAE. This is either the high frequencies (0.37 to 0.50 cyc/pixel) to evaluate error from image noise, or the entire range up to Nyquist frequency (0 to 0.5 cyc/pixel) to assess the average variation from the NS-SFR. The resulting MAEs were assigned to the coordinates for each system based on the edge parameters, which resulted in a 4D MAE coordinate system, sized 10x21x11x6 (angle, contrast, ROI height and radial distance).

The data is presented in two ways: plotting MAE for individual parameters (Figure 5.4) and MAE for multiple parameters (Figure 5.5).

The standard parameters (cf. § 2.4.1 & 4.2.4) were used as the fixed baseline, expanding a single parameter at a time (angle, contrast and then ROI height), covering MAE across that parameter range. Figure 5.4 contains plots of MAE against the tested parameters for the three LSF FWHM thresholds per camera system. Since noise introduces high-frequency error (cf. § 2.5.1), the high-frequency range was used when calculating the MAEs to demonstrate the highest variation from the parameters. Results indicate that these, now calibrated, NS-SFRs follow expected parameter characteristics. For each system, the system e-SFR estimate range error floor is consistent at an MAE of 0.1, potentially indicating the average natural scene image noise floor. Using more input NS-SFR data, for example, between DSLR 1 and DSLR 2 datasets, or using the 50th percentile threshold over 10th percentile, the variation around the error floor decreases.

The edge angle MAE range remains at the error floor up to approximately 35°, before the error increases. The edge contrast MAE range remains unchanging for the TIFF pipeline, fluctuating around the error floor. RAW data shows a low contrast MAE range increase not

present in the TIFF data, indicating non-linear demosaicing in the TIFF pipeline (cf. § 4.3.2). For DSLR 1, ROI height is stable across the entire range, whilst the DSLR 2 data has increased variation but still demonstrates an average trend of 0.1 MAE.

There are discrepancies when comparing this data, Figure 5.4, with DSLR 1 TIFF test chart counterpart (Figure 4.10). Unlike the previous study (Figure 4.10), the characteristic angle variation (cf. § 2.4.1) is depicted in Figure 5.4. The non-linear high contrast error seen in the test chart data is not present in this natural scene counterpart. This suggests that the previous speculation of the anisotropic nature of the sensor array may result in the demosaicing interpolation having spatial effects (cf. § 4.3.2) to be an artifact of highly isolated edges, i.e. test charts. Step-edges are surrounded by scene textures, noise, and other scene artifacts within natural scenes. This results in edges that are not as easily enhanced in the demosaicing process. Further work is required to confirm this hypothesis.

Figure 5.5 depicts a matrix illustrating MAE introduced in the NS-SFR data across multiple parameters. This figure uses the MAE from the TIFF DSLR 1 NS-SFR data, calculated across the entire spatial frequency range up to Nyquist frequency (0 to 0.5 cyc/pixel). Similar results to the parameter variation study (cf. § 4.3) are observed here; notably, parameters prone to noise corruption compile to increase MAE.

Overall, results validate that the selected parameter ranges are stable and suitable for system e-SFR estimation. Furthermore, since images used for this validation originate from systems with minimal ISP, the plots also indicate that the edge contrast range could be increased, including more edges extracted from the natural scenes, potentially improving system e-SFR precision. That said, the contrast range is kept as the standard, 0.55 to 0.65 Michelson contrast, to diminish variance introduced by non-linear sharpening when implementing the method on non-linear systems, as stated in Chapter 4 (cf. § 4.4).

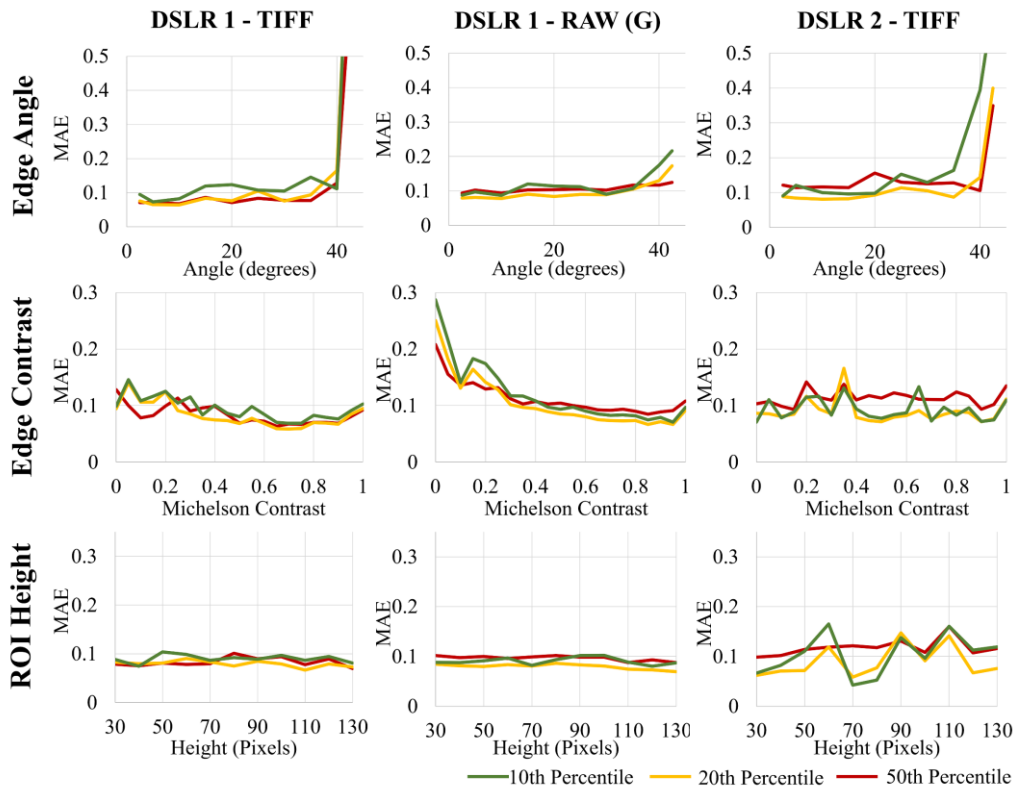


Figure 5.4 The comparison of MAE introduced by e-SFR parameters, edge angle, edge contrast and ROI height, between spatial frequencies 0.37 to 0.50 cyc/pixel, from NS-SFRs calculated using three LSF FWHM thresholds for three systems [124].

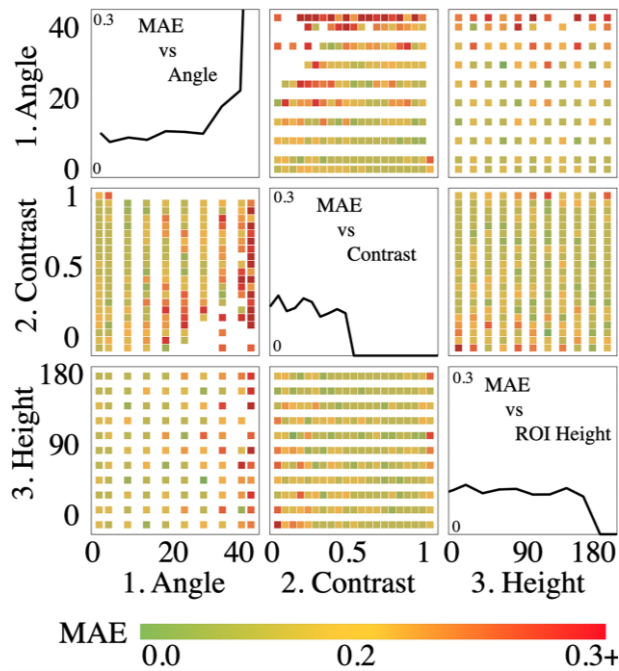


Figure 5.5 Matrix of variation introduced to the NS-SFR measurement when changing edge angle, edge contrast, ROI height and ROI width, across spatial frequencies 0.0 to 0.5 cyc/pixel. The diagonal of the matrix plots MAE introduced by each SFR parameter. Data is from the DSLR 1 system using a TIFF pipeline.

5.3 System e-SFR Estimation Results

5.3.1 Characterised Systems and the Respective Image Datasets

Three digital camera systems were used to capture a large number of scenes used for producing three image datasets. These datasets are examples of different system pipelines, but not necessarily examples of different types of applications. This section details how each dataset was used in its entirety to estimate the corresponding system e-SFR. The camera and settings were as follows:

1. DSLR 1 – *Nikon D800*, 36.3-megapixel sensor ($4.87\mu\text{m}$ pixel pitch), with a *Nikon AF-S 24-70mm f/2.8G IF-ED* lens at 24mm $f/4$ (wide-angle).
2. DSLR 2 – *Nikon D800*, 36.3-megapixel ($4.87\mu\text{m}$ pixel pitch) sensor, with a *Carl Zeiss 135mm f/2 Apo-Sonnar T** lens at $f/4$ (telephoto).
3. Smartphone camera – *Apple iPhone7*, 12-megapixel sensor ($1.22\mu\text{m}$ pixel pitch), with a 3.99mm lens at $f/1.8$.

The DSLR 1 camera dataset contains 1866 images. The wide-angle focal length of this system allows for wide depth of fields, thus reducing the number of out of focus edges. It is worth noting that wide-angle lenses tend to have a wide variation in performance across the imaging circle. This allows the proposed approach to be assessed against significant optical e-SFR variations across the field of view. Two versions of this dataset were produced: the first comprises of uncompressed TIFF image files without sharpening or denoising applied in the demosaicing process; the second is comprised of the green channel of the RAW (mosaiced) sensor image.

The DSLR 2 camera dataset contains 1009 images. Although the camera body is identical, the two DSLR systems have different optical system performance characteristics due to their lens designs. In contrast to wide-angle lenses, telephoto optics tend to have a smaller performance variation across the frame. Telephoto lenses provide a shallower depth of field, producing many edges out-of-focus, which are not suitable for accurate system e-SFR estimations. This second dataset tests the estimation method's robustness against tighter optical performance tolerances for each radial segment and its effectiveness in eliminating defocused edges from the calculation. Again, only uncompressed TIFF files were used with the denoising and sharpening turned off during demosaicing.

The smartphone camera dataset, containing 2008 images, was set to test the suitability of the proposed method for assessing the performance of non-linear pipelines. Smartphone camera images are subject to heavy non-linear ISP and JPEG compression. Both processes are scene dependent and introduce different artifacts to the captured edges. Since many modern systems use non-linear ISP, studying how the proposed method behaves with such a system is essential.

During capture, the shutter speed and ISO gain settings were set to best expose the images (determined by the camera's global exposure meter). Most images were captured using the camera systems' base ISO gain. As a result, each dataset contained a range of noise levels and thus different SNR distributions. Figure 5.6 illustrates the distribution of the ISO gain settings, as reported in the EXIF image metadata.

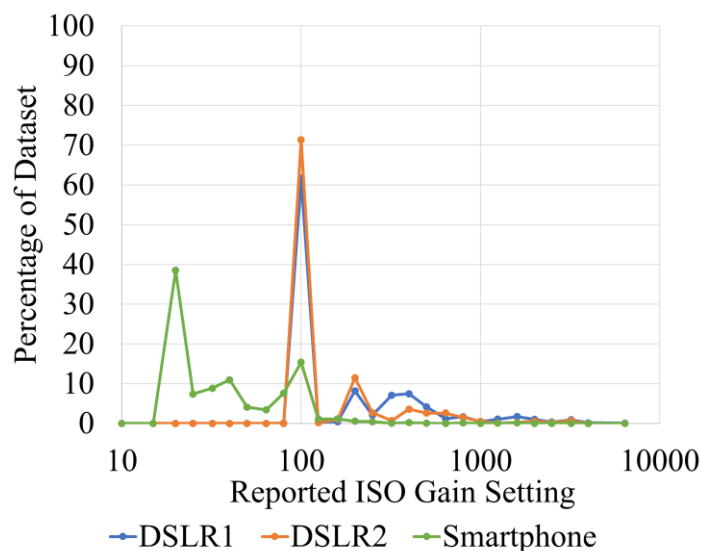


Figure 5.6 The distribution of the ISO gain settings for the images in the three datasets [146].

Each of the three datasets is comprised of images that were subsequently classified according to three different scene locations: man-made exteriors, indoor scenes, and nature scenes. The AlexNet CNN [147], after being transfer learned [148], was employed for the classification. This classification process is detailed in Chapter 6 (cf. § 6.2). Figure 5.7 shows the resulting scene class distributions for each dataset.

The camera systems were characterised using the ISO12233:2017 slanted-edge method [5] to obtain e-SFRs across the camera frame. The frame was divided into six radial distances in the same fashion as in the e-SFR estimates from natural scenes (cf. § 5.1.4). The test chart edges were positioned so that they were evenly distributed around each radial segment. The

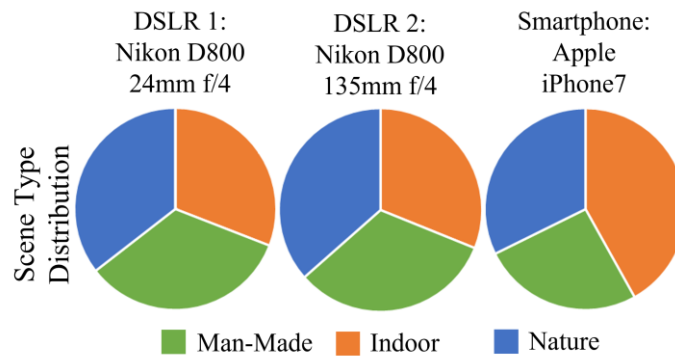


Figure 5.7 Scene type distributions for the three datasets, classified by re-trained AlexNet. Adopted from [146].

e-SFRs within these six segments were averaged using the same weighted values as the system e-SFR estimation [74] (cf. § 5.1.4). These e-SFRs were calculated from both the TIFF files (with sharpening and denoising turned off in the demosaicing process) and the green channel of the RAW files. The difference in these e-SFR results is negligible, as shown in Appendix A. Hence only the TIFF ISO12233 e-SFRs were used as the system characterised data.

These averaged ISO12233 e-SFRs are considered the measurement against which estimates are assessed. For each average ISO12233 e-SFR measure, ± 1 standard deviation was calculated to provide a range deemed acceptable for assessing the accuracy of the estimates. One standard deviation was chosen as it provides the distribution where 68% of the test chart e-SFRs should be statistically found. An accurate system e-SFR estimate comparable to the ISO12233 e-SFR should have a mean within this region.

It is important to note that the radial segments are based on the imaging circle, so the area of the field of view segments are uneven. This results in a different number of edges used in the standard deviation calculations per segment. Table 5.1 provides the number of edges per radial segment for each system. Areas with fewer edges have increased uncertainty in the standard deviation, particularly in radial distances with greater performance variation, such as the corners of the field of view (cf. § 2.4.3). That said, the even distribution across each radial segment would provide the expected performance variation of each segment.

Radial Distance	Camera System											
	DSLR 1				DSLR 2				Smartphone			
	V	H	S	T	V	H	S	T	V	H	S	T
1	14	14	16	16	15	16	29	16	16	16	26	17
2	30	28	27	37	33	26	29	45	32	28	33	40
3	43	42	56	93	42	47	53	89	42	46	48	96
4	41	46	51	91	55	60	48	87	44	46	51	91
5	31	24	16	44	50	52	13	31	29	27	15	22
6	5	4	5	6	17	15	4	6	5	5	6	8

Table 5.1 Number of test chart edges per radial distance segment for each of the system e-SFR orientations (Vertical (V), Horizontal (H), Sagittal (S) and Tangential (T)).

5.3.2 Region of Interest Utilisation

The system e-SFR estimation is based on a highly selective process. The process yields small numbers of suitable natural scene step-edges. The following sections detail the yield of NS-SFR that are used in the system e-SFR estimations when using vertical and horizontal (cf. § 5.3.2.1) and sagittal and tangential (cf. § 5.3.2.2) edge orientations.

5.3.2.1 Vertical and Horizontal Edge Orientations

The number of isolated vertical and horizontal ROIs depends on the system pixel resolution as well as ISP. For example, an average of 19 ROIs per image were selected from the smartphone camera system, whilst for the higher resolution DSLR 1 and DSLR 2 TIFF datasets, there were 72 ROIs per image. Furthermore, many of these ROIs were deselected in the process of thresholding the most likely suitable step-edge inputs (cf. § 5.1.3) and setting limits to the edge parameter ranges (cf. § 5.1.4). When using the 10th percentile of the LSF FWHM distribution, on average, only 3.41% of isolated ROIs were utilised in the system e-SFR estimation, corresponding to 1.69 ROIs per image. As the percentile threshold increases to include more NS-SFR data, utilisation increases linearly, with the 20th percentile utilising 6.50% of the ROIs (3.23 natural scene ROIs per image) and 50th percentile 16.83% of the ROIs (8.31 natural scene ROIs per image).

Figure 5.8 plots the number of ROIs extracted from the DSLR 1 TIFF dataset per radial segment and the number of these ROIs utilised in the vertical system e-SFR estimation, derived using the three LSF FWHM thresholds. The distributions show approximately 80% fewer ROIs in the centre (1/6) and corners (6/6) of the frame than the partway segments. The two other systems, as well as the horizontal NS-SFRs, exhibit the same trend.

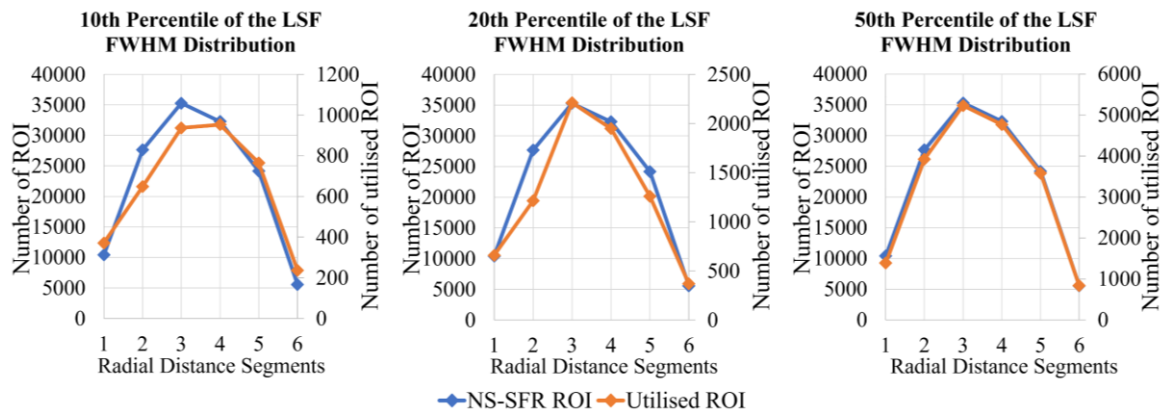


Figure 5.8 Number of ROIs isolated and the number of utilised ROIs per radial distance for the vertical e-SFR estimation from TIFF DSLR 1 dataset for each LSF FWHM distribution threshold.

Table 5.2 provides the utilisation percentages for the radial segments centre (1/6), partway (3/6) and corners (6/6), as well as the entire frame. The ‘average frame utilisation’ is the mean utilisation of the entire frame across all the datasets. Three sets of percentages are presented in this table, showing the ROIs used for each percentile threshold.

Knowing that percentile thresholds would initially isolate approximately 10%/20%/50% of NS-SFR data, it can be deduced that, on average, parameter range thresholds further deselect 66% of data. The most significant parameter constraint is the narrow edge contrast range of 0.55 to 0.65 Michelson contrast.

System	Radial Distance	Total 10th Percentile Utilisation (%)	Total 20th Percentile Utilisation (%)	Total 50th Percentile Utilisation (%)	
DSLR 1	TIFF Dataset	1/6	3.57	6.33	13.34
		3/6	2.66	6.27	14.82
		6/6	4.25	6.63	15.03
		Frame	2.89	5.66	14.57
	RAW Dataset	1/6	5.07	9.30	18.04
		3/6	3.99	8.13	17.12
	6/6	4.36	6.18	16.33	
	Frame	4.37	7.19	16.69	
DSLR 2	TIFF Dataset	1/6	2.68	5.82	13.07
		3/6	2.15	4.55	13.76
		6/6	2.72	5.85	14.85
		Frame	2.33	5.00	13.81
Smartphone	TIFF Dataset	1/6	5.18	6.92	19.76
		3/6	3.34	6.72	19.80
		6/6	4.95	11.42	28.18
		Frame	4.04	8.16	22.24
Average Frame Utilisation			3.41	6.50	16.83

Table 5.2 Percentage utilisation for each dataset used to estimate the vertical system e-SFR for radial distance segments representing the centre, partway, corners, and entire frame.

5.3.2.2 Sagittal and Tangential Orientations

Converting vertical and horizontal edges to sagittal and tangential orientations (cf. § 2.6.2 & 5.1) further reduces the data available for the e-SFR estimations. Unlike vertical and horizontal, sagittal and tangential orientations have a specific position in the imaging circle. From the DSLR 1 TIFF dataset, 261,380 vertical and horizontal ROIs were isolated; this number is reduced by 71% when converting the orientation, resulting in only 76,299 sagittal and tangential ROIs in total.

The deselection processes of the LSF FWHM thresholding and parameter restrictions follow the same percentage decreases as for the vertical and horizontal data (cf. § 5.3.2.1). Figure 5.9 compares the number of isolated sagittal ROIs and the utilised ROIs, for the three LSF FWHM thresholds. Table 5.3 contains sagittal ROI utilisation as a percentage for each dataset.

Sagittal orientation has a slightly greater decrease in ROIs in the centre and corners of the frame than the vertical ROIs, with approximately 85% fewer edges than partway regions. This data confirms that, on average, parameter thresholds deselect a further 66% of the data after LSF FWHM selection.

The abovementioned trends are consistent across all datasets for both sagittal and tangential orientations.

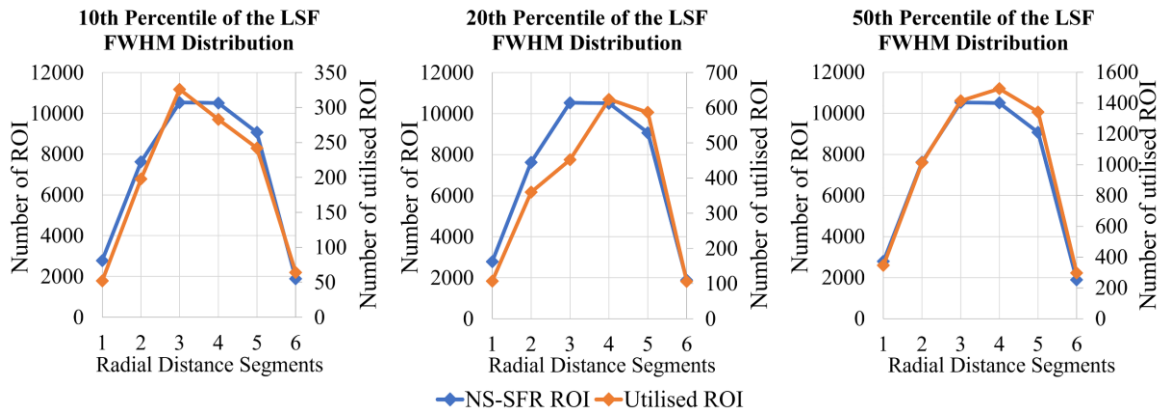


Figure 5.9 Number of ROIs isolated and the number of utilised ROIs per radial distance for the sagittal e-SFR estimation from TIFF DSLR 1 dataset for each LSF FWHM distribution threshold.

System	Radial Distance	Total 10th Percentile Utilisation (%)	Total 20th Percentile Utilisation (%)	Total 50th Percentile Utilisation (%)
DSLR 1	1/6	1.87	3.84	12.43
	TIFF	3.10	4.29	13.42
	Dataset	3.39	5.62	15.69
	Frame	2.75	5.27	13.92
	1/6	5.51	9.83	16.73
	RAW	4.63	8.56	17.34
	Dataset	6.46	6.46	17.05
	Frame	4.13	7.71	16.33
DSLR 2	1/6	3.40	6.45	15.28
	TIFF	2.52	5.07	13.40
	Dataset	2.32	5.05	14.46
	Frame	2.33	5.03	13.74
Smartphone	1/6	4.79	6.46	17.71
	TIFF	2.86	8.26	21.36
	Dataset	6.17	11.52	29.22
	Frame	3.92	8.27	22.21
Average Frame Utilisation		3.28	6.57	16.55

Table 5.3 Percentage utilisation for each dataset used to estimate the sagittal system e-SFR for radial distance segments representing the centre, partway, corners, and entire frame.

5.3.3 DSLR (Near-Linear) Camera System e-SFR Estimations

Absolute error at 25%, 50% and 75% of the Nyquist frequency were used to illustrate the accuracy of the vertical estimates from both DSLR systems in Figure 5.10. Nyquist frequency metrics were used, rather than the MTF50/20/10 metrics (cf. § 2.6.1) with higher standard deviations. Horizontal estimates show similar trends to the vertical and therefore are not included in this chapter. Additional system e-SFR metrics from all tested datasets can be found in Appendix B.

The Standard Error (SE) is used for the error bars in Figure 5.10, indicating the accuracy of the mean system e-SFR estimates in relation to the number of utilised ROIs, N , and the resultant standard deviation, σ . SE is expressed as:

$$SE = \frac{\sigma}{\sqrt{N}} \quad (5.1)$$

SEs indicate that the precision is high for most system e-SFR estimates; the SE decreases with the increase of the LSF FWHM threshold. As mentioned (cf. § 5.3.2), the number of utilised ROIs scales linearly with the LSF FWHM threshold, thus going from the 10th to 50th

percentile, on average, increases the number of ROIs by five times, whilst the increase in the standard deviation is only 1.1 times, thus improving the SE. The radial distance segments with large SE, for example, 6/6 for the DSLR 2 dataset, are due to the small frame area having fewer ROIs, in combination with a greater optical performance variation (cf. § 2.4.3) causing higher standard deviations.

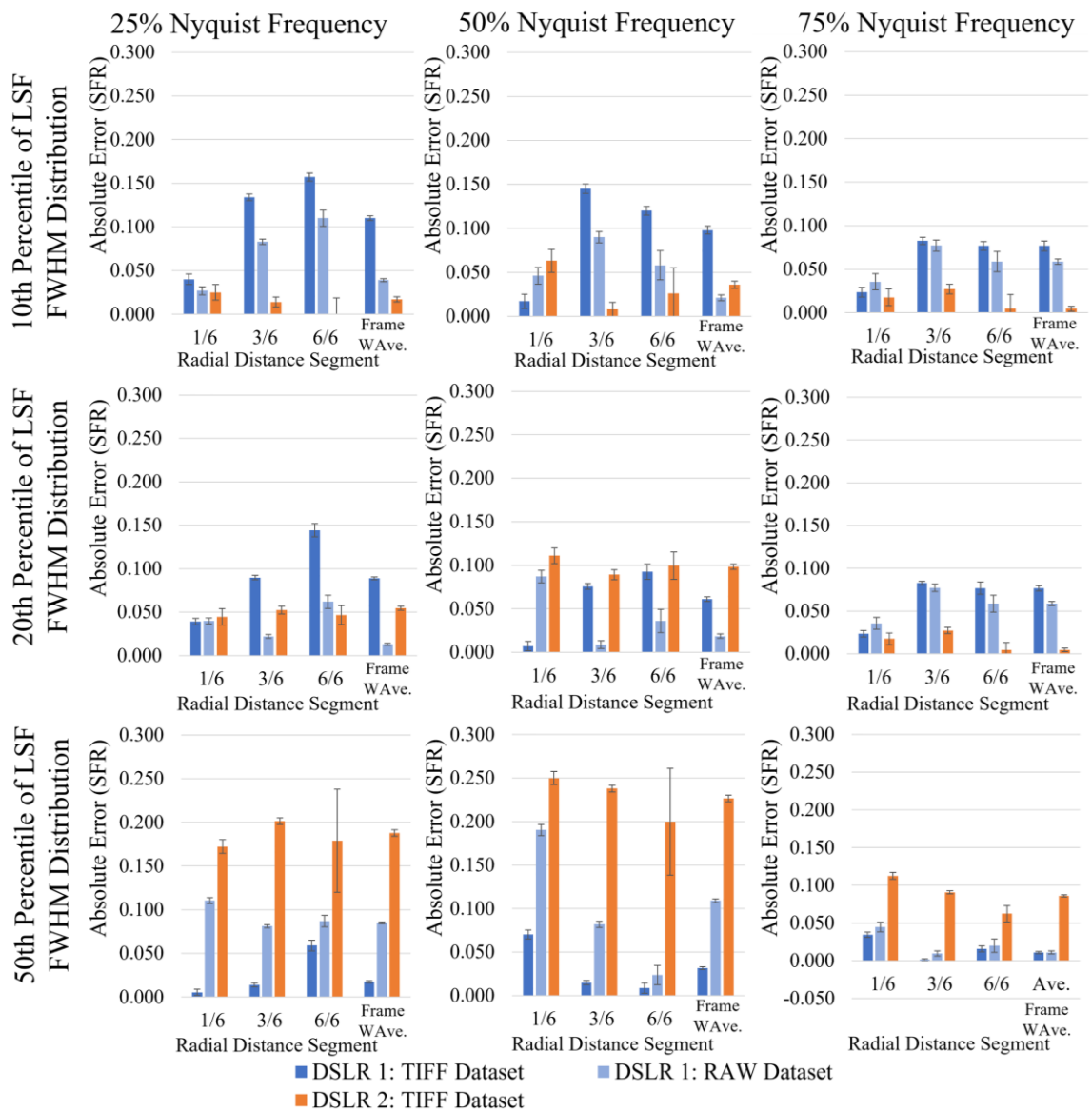


Figure 5.10 Absolute error at the 25%, 50% and 75% of the system vertical e-SFR estimate Nyquist frequency, derived from three LSF FWHM thresholds, for the DSLR systems at radial segments 1/6, 3/6 and 6/6 as well as the weighted average (WAve.) of the entire frame. Adopted and expanded on from [146].

Figures 5.11 and 5.12 contain vertical system e-SFRs derived from DSLR 1 and DSLR 2 datasets, respectively. The 10th percentile of the LSF FWHM distribution was first used in the calculations to begin the analysis of the accuracy of the estimation. The first column shows the estimated system e-SFR in relation to the equivalent ISO12233 e-SFR radial segment mean and ± 1 standard deviation, i.e. e-SFR envelope, indicating the acceptable error range (cf. § 5.3.1). The second column shows the absolute error of the estimations in respect to the ISO12233 mean e-SFR and its standard deviation envelope. The third column depicts radial distance segments from which data is measured; three radial distance segments and the weighted average are plotted to provide an overview of estimates across the frame.

These plots combined with data presented in Figure 5.10 indicate moderate to high accuracy in system e-SFR estimations calculated using the 10th percentile threshold. Excluding frequencies above 75% Nyquist, estimations are within 0.15 SFR of the ISO12233 e-SFR. RAW DSLR 1 dataset returned a more accurate system e-SFR estimate than the TIFF counterpart. High frequencies were overestimated throughout all RAW system e-SFR estimates, a bias associated with high noise levels [39]. In contrast, DSLR 1 TIFF results do not show this noise bias, indicating that the previously discussed incorporation of non-linear demosaicing is present in the TIFF pipeline (cf. § 4.3.2). Also, it is observed that the results from the TIFF files, although following the same trend as the ISO12233 e-SFRs, are overestimated throughout the range of spatial frequencies.

The DSLR 2 system e-SFR estimates that were calculated using the 10th percentile of the LSF FWHM distribution, with fewer pictorial natural scenes, has higher accuracy than from the DSLR 1, staying within 0.06 SFR of the standard e-SFR. This suggests that the proposed method is robust against shallow depth of field and is proficient with obtaining telephoto lens performance with tight precision tolerances.

Finally, the weighted average estimates use all isolated ROIs across the frame. For both systems, the resultant e-SFRs correspond to the average ISO12233 e-SFR, staying within one standard deviation. They deliver the most consistent estimate of the system e-SFR from natural scene inputs.

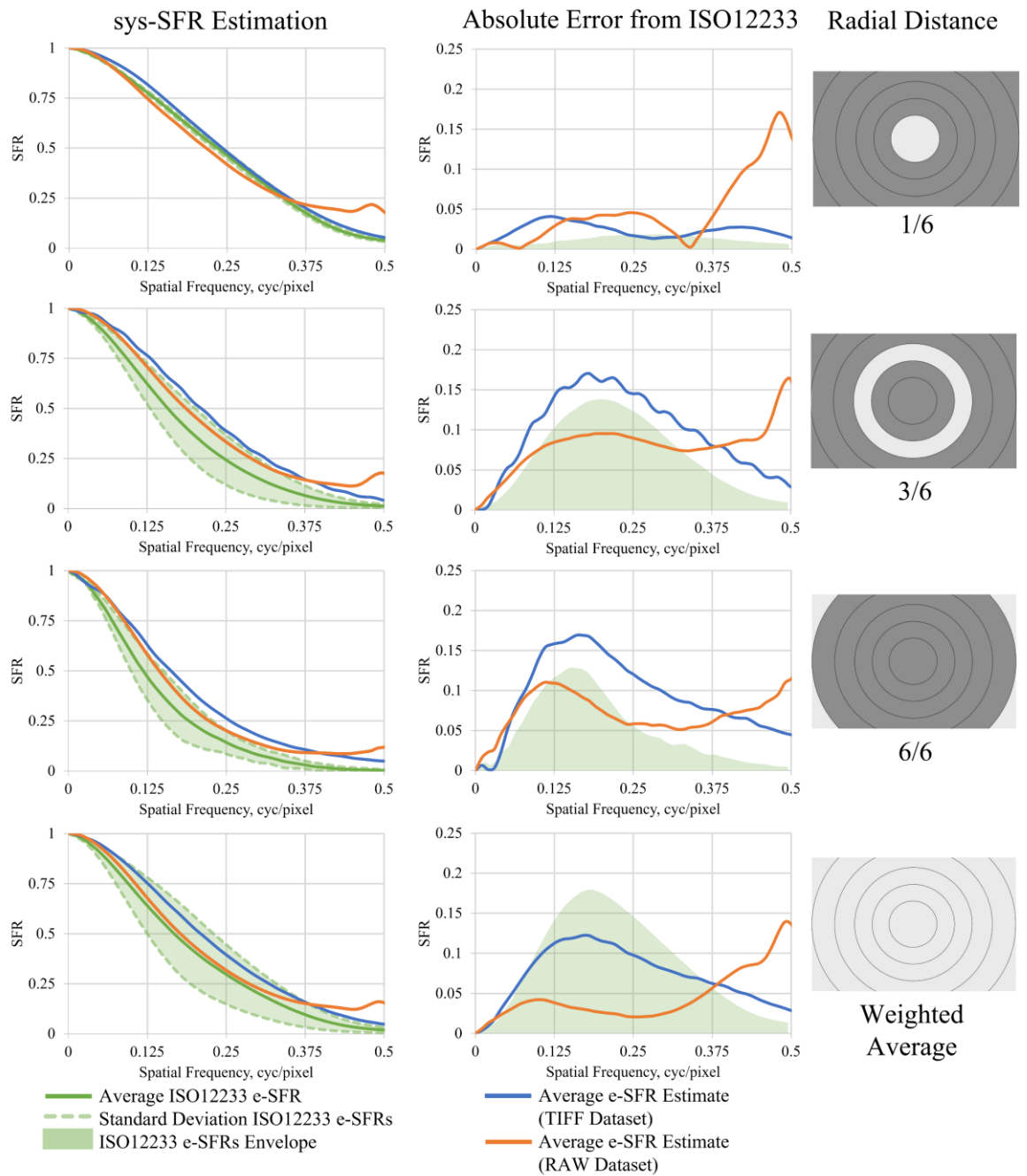


Figure 5.11 DSLR 1 vertical system e-SFR estimation for three radial distances out of six and a weighted mean of all six radial distances. The first column contains the estimated system e-SFR in relation to the ISO12233 e-SFR. The second column contains the absolute error between the estimated system e-SFR from the mean of the ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [145, 146].

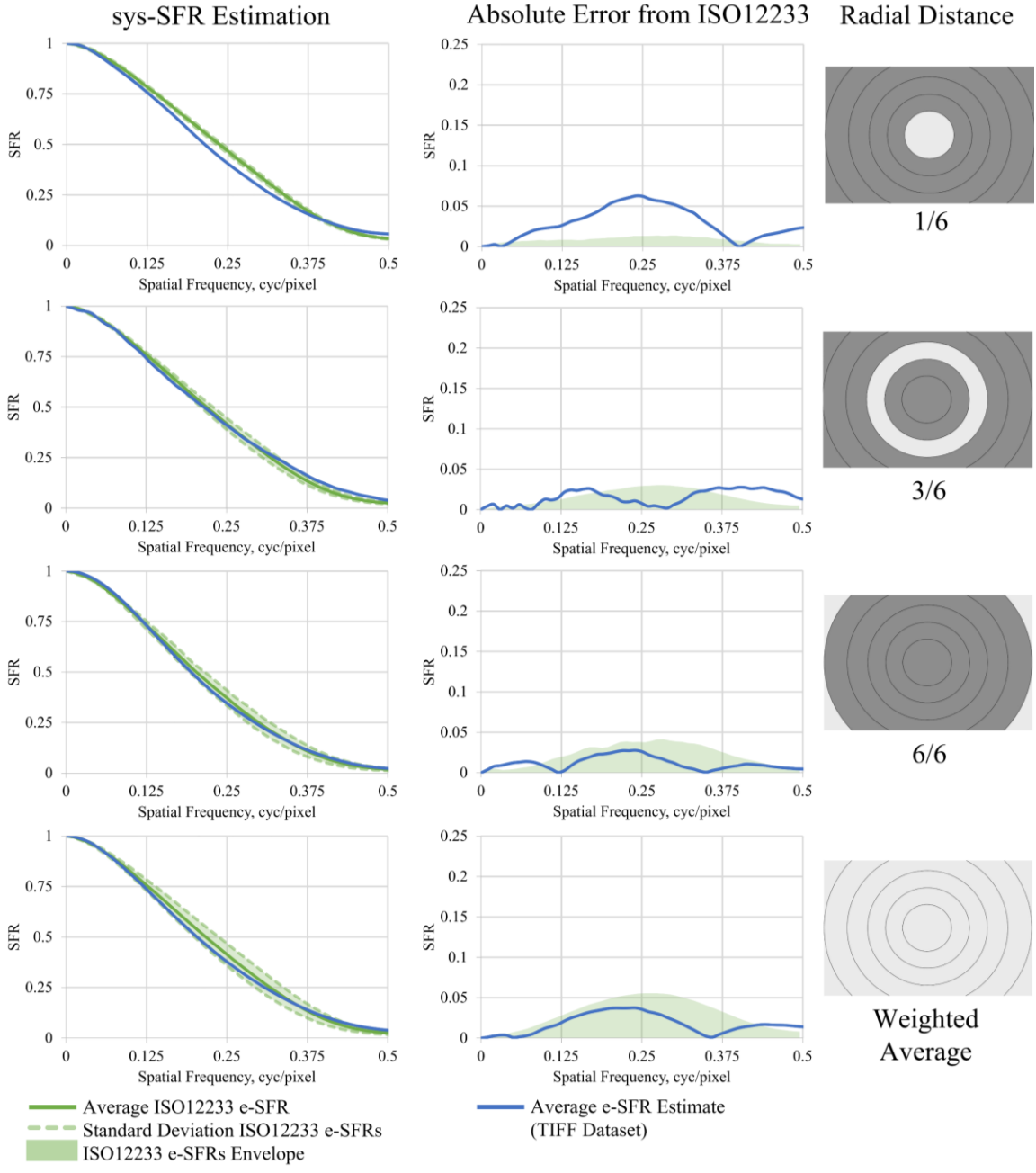


Figure 5.12 DSLR 2 vertical system e-SFR estimation for three radial distances out of six and a weighted mean of all six radial distances. The first column contains the estimated system e-SFR in relation to the ISO12233 e-SFR. The second column contains the absolute error between the estimated system e-SFR from the mean ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [146].

Estimating the system e-SFR using wider LSFs incorporates more of the NS-SFR data. Consequently, the inclusion of wider LSFs resulted in e-SFR estimations depicting lower performance. Figures 5.13 and 5.14 compare the weighted average system e-SFR estimates for all three LSF FWHM distribution thresholds, the 10th, 20th and 50th percentiles, for DSLR 1 and DSLR 2 datasets, respectively. As mentioned (cf. § 5.1), this additional data increases the probability of introducing unsuitable step-edge inputs.

The estimates from the DSLR 1 stay within one standard deviation of the ISO12233 e-SFR for all three thresholds. That said, the e-SFRs derived from RAW files become underestimated when increasing the threshold, whilst those derived from the TIFF counterpart increase in accuracy. This is due to lower performance edges offsetting the edges that produce overestimation. Whilst, e-SFRs derived from the DSLR 2 dataset become underestimated, providing results below the limits of one ISO12233 e-SFR standard deviation.

Given the above results, the 10th percentile is recommended, as it increases the probability of selecting the suitable step-edge inputs for an estimate of the system e-SFR. In addition, this threshold is shown to provide accurate results for both DSLR systems. If the NS-SFR data is limited, this threshold can be increased to include more of the selected data. For an optical system with extensive performance variation across its imaging circle, such as a wide-angle lens (DSLR 1), there is a greater overlap of system performance and scene content. Thus, accuracy limits are more forgiving, allowing this LSF FWHM threshold to include more data without decreasing accuracy, although this may decrease the estimation precision. In contrast, a telephoto lens (DSLR 2) with narrow performance tolerances is not suited to wider LSF FWHMs, as it leads to underestimated e-SFRs.

System e-SFR estimates for all six radial distance segments and the weighted average are included in Appendix C. This data is presented three times per system, each computed using a different LSF FWHM distribution threshold (10th, 20th and 50th percentile).

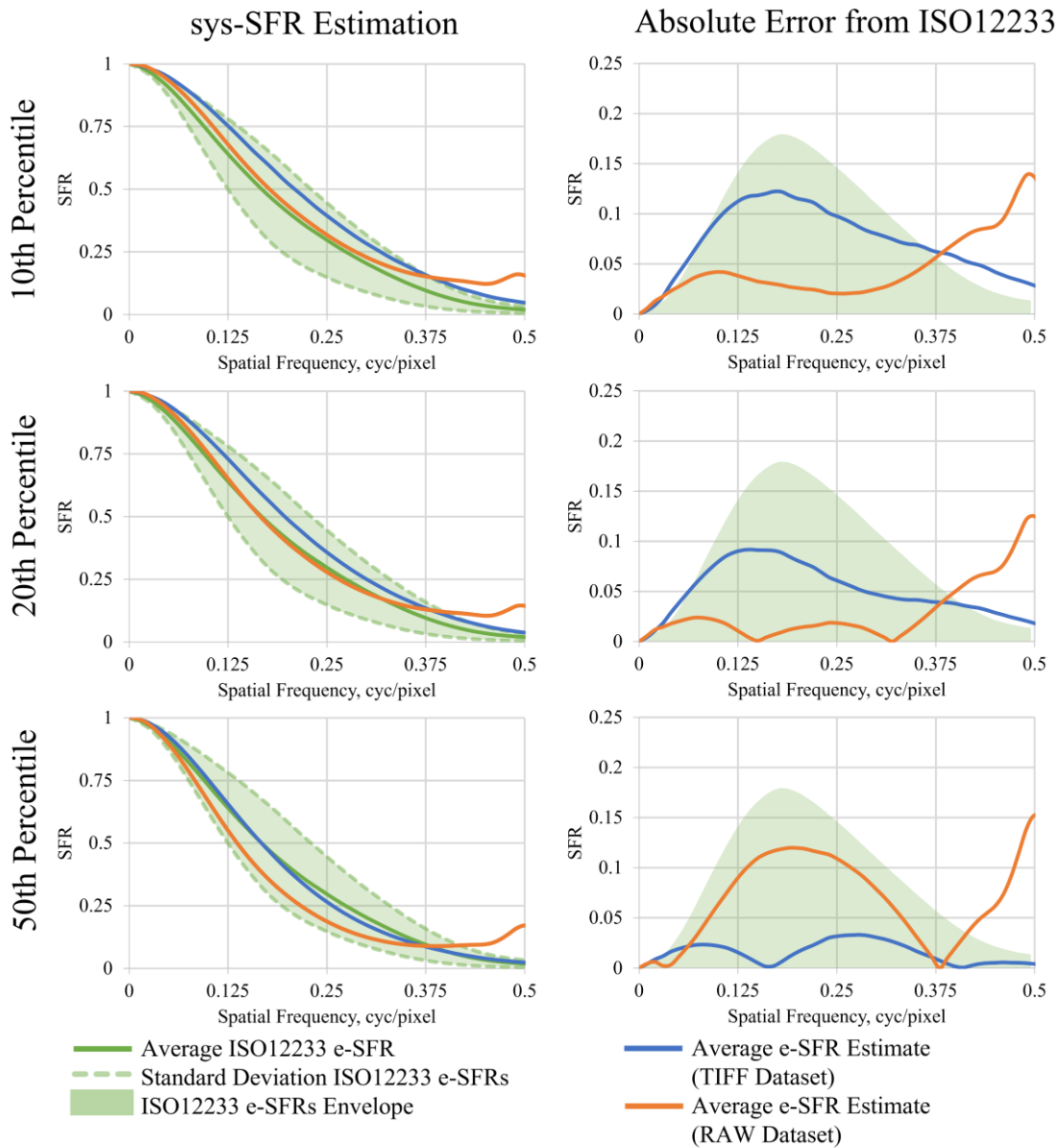


Figure 5.13 Comparison between the DSLR 1 (TIFF and RAW datasets) vertical weighted average system e-SFR estimates derived from three LSF FWHM distribution thresholds, 10th, 20th and 50th percentiles.

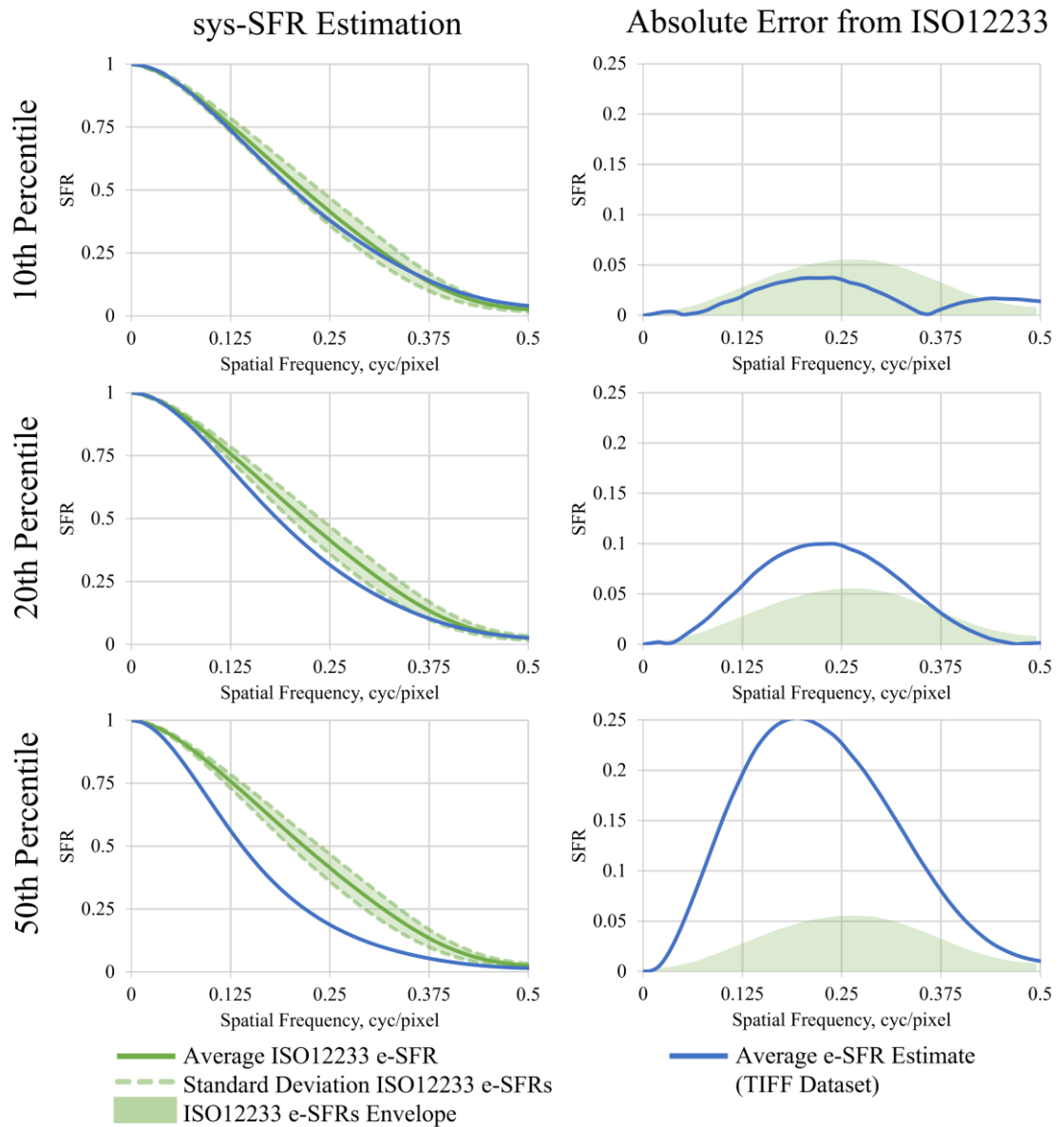


Figure 5.14 Comparison between the DSLR 2 vertical weighted average system e-SFR estimates derived from three LSF FWHM distribution thresholds, 10th, 20th and 50th percentiles.

The DSLR 1 RAW image dataset was segmented into three scene classes. These were man-made (exterior), indoor and nature scenes, as defined using the transfer learned AlexNet CNN detailed in Chapter 6 (cf. § 6.1.2). Implementing these segmented datasets to the methodology provide close approximations to the test chart e-SFR measurement between 0 cyc/pixel and 75% of the Nyquist frequency, as shown in Figure 5.15. These results indicate that a test chart e-SFR can be derived from various scene classes when using a linear system.

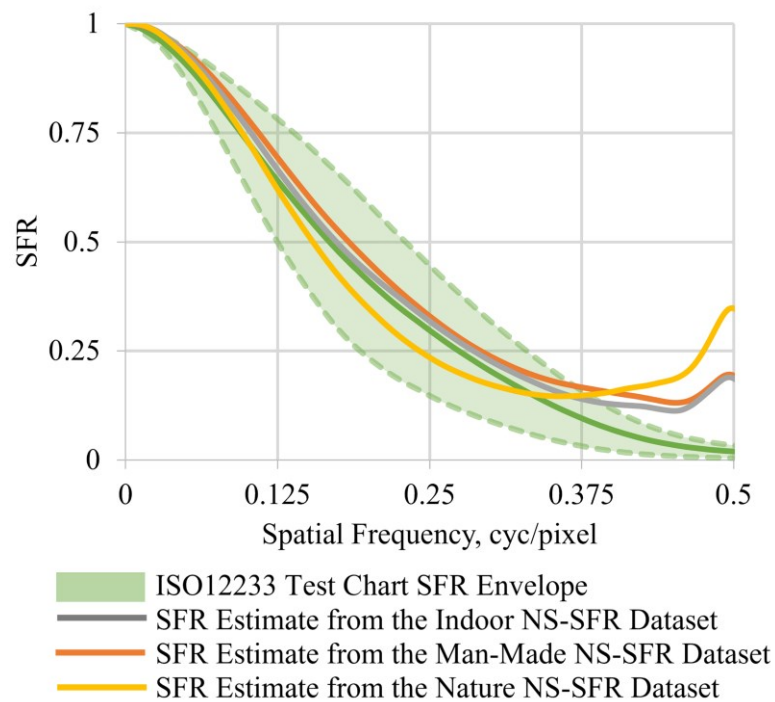


Figure 5.15 Weighted average vertical system e-SFR estimate from the DSLR 1 sub-datasets of indoor, man-made and nature scene classes. These estimates are compared to the average ISO12233 slanted edge e-SFR and its standard deviation envelope.

Figures 5.16 and 5.17 illustrate the estimated DSLR 1 and DSLR 2 sagittal system e-SFRs, derived from converting vertical and horizontal NS-SFR edge orientations (cf. § 5.1.2). System e-SFR estimates at 25%, 50% and 75% Nyquist frequency are plotted against radial distance. The estimates from the tangential orientation are found in Appendix D.

Through converting vertical and horizontal edges to sagittal and tangential, 71% of isolated NS-SFR data is deselected because they do not conform to the desired orientation (cf. § 5.3.2.2). With this smaller quantity of data available, the accuracy of the e-SFR estimate decreases. Both DSLR systems have overestimated sagittal system e-SFRs when using the 10th percentile of the LSF FWHM distribution. As with vertical e-SFRs, the wider LSF FWHMs decreased estimated performance. Further, the proposed method once again is

shown to work effectively with an optical system with tighter performance tolerances, i.e. the DSLR 2 telephoto lens. The DSLR 1 dataset requires a 50th percentile threshold to provide the most accurate sagittal estimates for RAW and TIFF pipelines, whilst DSLR 2 dataset delivers the greatest accuracy at the 20th percentile. This discrepancy between the two systems indicates that more edges are required when estimating sagittal or tangential e-SFRs, particularly for lenses with wide performance variation. Expanding the LSF FWHM distribution threshold is not the most appropriate method to achieve this, as it introduces edges with unknown modulation content; measuring NS-SFRs from larger datasets is more suitable.

The above indicated that the proposed method to estimate system e-SFR is best suited for measuring vertical and horizontal e-SFRs, since these edge orientations commonly occur in natural scenes, unlike sagittal and tangential orientated edges. That said, sagittal and tangential system e-SFRs can be estimated with reasonable accuracy regarding the ISO12233 test chart output when the system is linear, but the caveats discussed must be considered when implementing it.

When implementing the e-SFR estimation, several adjustments can be made to the method for specific application requirements. The number of radial distances may be increased to deliver a higher resolution measurement across the frame. The method can be modified to map sagittal and tangential e-SFRs to 2D coordinates, representing performance across the field of view. These adjustments require extensive dataset inputs as they introduce further segmentation of the isolated NS-SFR data.

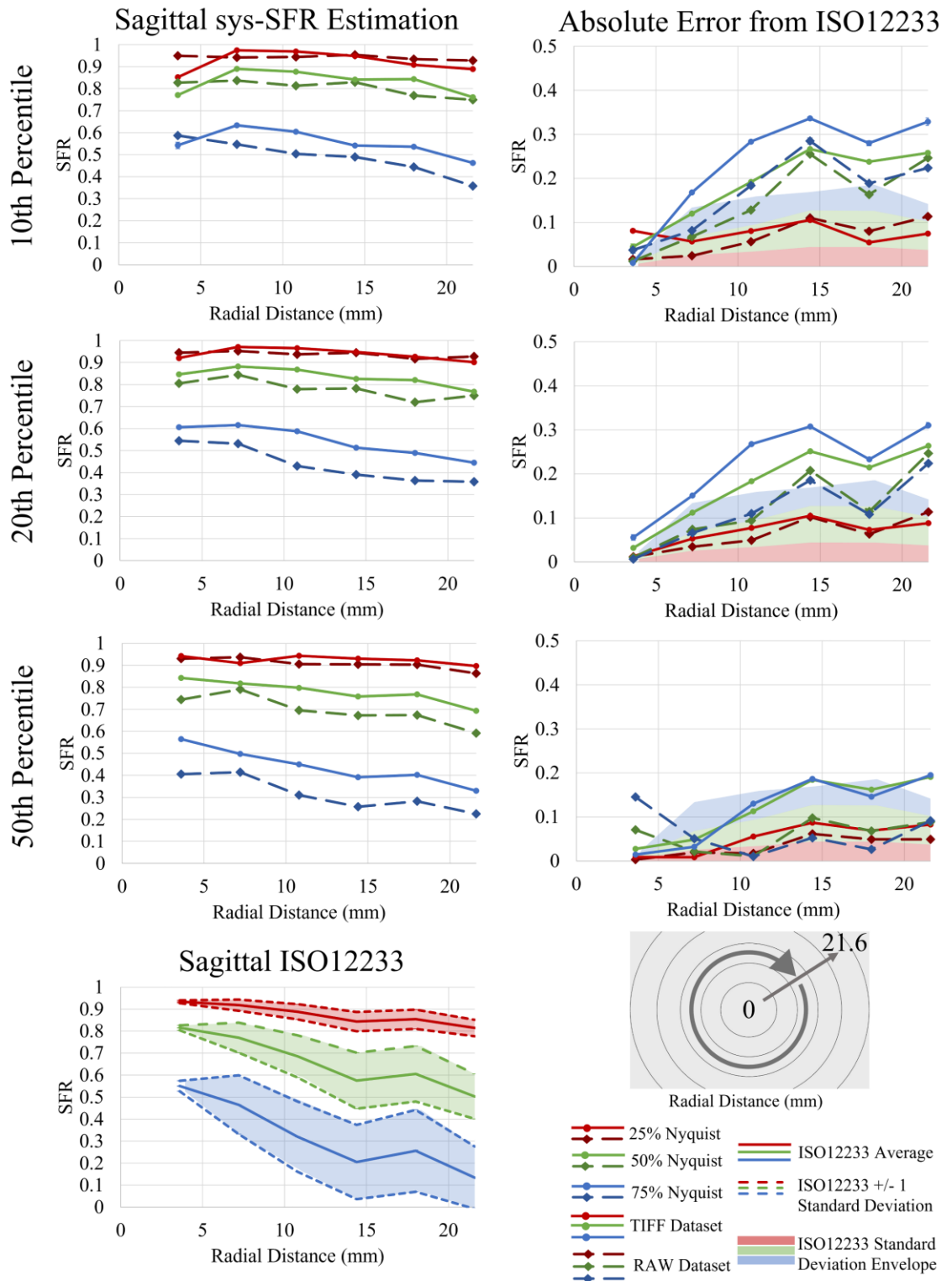


Figure 5.16 25%, 50% and 75% Nyquist frequency of DSLR 1 sagittal TIFF and RAW system e-SFR estimates are plotted against the radial distance. Estimates are derived from the 10th, 20th and 50th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO1223 e-SFR is also plotted, with a standard deviation of the ISO1223 shaded.

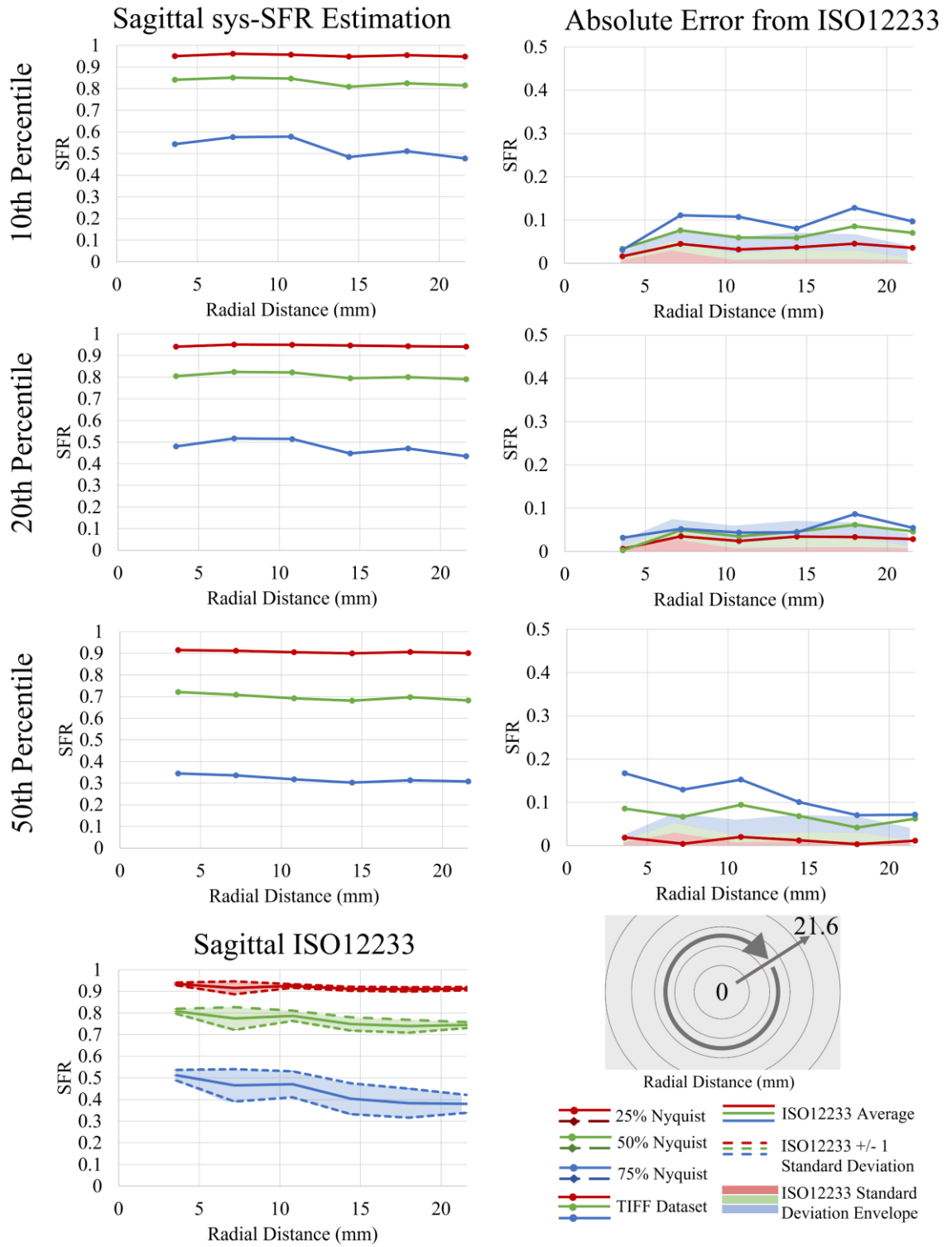


Figure 5.17 25%, 50% and 75% Nyquist frequency of DSLR 2 system e-SFR estimates are plotted against the radial distance. Estimates are derived from the 10th, 20th and 50th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO12233 e-SFR is also plotted, with a standard deviation of the ISO12233 shaded.

5.3.4 Smartphone (Highly Non-Linear) Camera System

The absolute error for the vertical system e-SFR estimate at the 25%, 50% and 75% Nyquist frequency were measured for the non-linear smartphone dataset, as shown in Figure 5.18. These errors are greater than those calculated from performance estimations from the linear systems, as shown in Figure 5.10. The error increases when using a wider LSF in the estimation. The most accurate results are from the weighted average system e-SFR estimation, computed using the 10th percentile of the LSF FWHM distribution, staying within 0.10 for the 25% Nyquist frequency and 0.05 for the 50% and 75% Nyquist frequency.

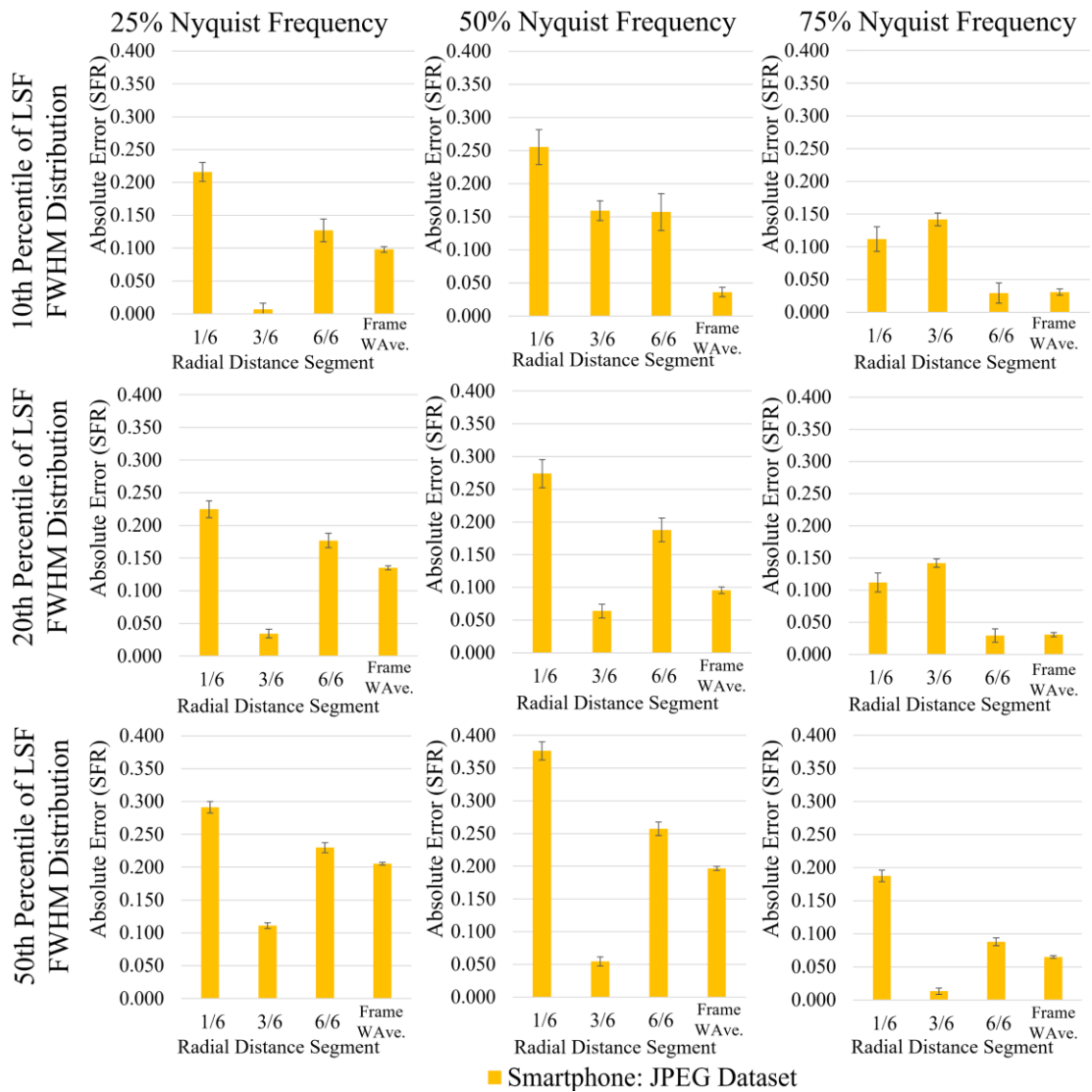


Figure 5.18 Absolute error at the 25%, 50% and 75% of the smartphone vertical e-SFR estimate Nyquist frequency, derived from three LSF FWHM thresholds, at radial segments 1/6, 3/6 and 6/6 as well as the weighted average (WAVE.) of the entire frame. Adopted and expanded on from [146].

The smartphone system e-SFR estimations for radial distance segments 1/6, 3/6 and 6/6, as well as the frame weighted average, computed using the 10th percentile of the LSF FWHM distribution, are plotted in Figure 5.19. Also displayed in this figure are the corresponding ISO12233 e-SFRs and standard deviation envelopes. Figure 5.20 depicts the smartphone weighted average estimate in relation to both the texture-MTF (i.e. dead leaves MTF), measured using the Imatest spilled-coins test chart [32], and the ISO12233 e-SFR. The texture-MTF is designed to simulate an average natural scene signal, resulting in a more faithful camera response for systems subjected to non-linear processing (cf. § 1.1).

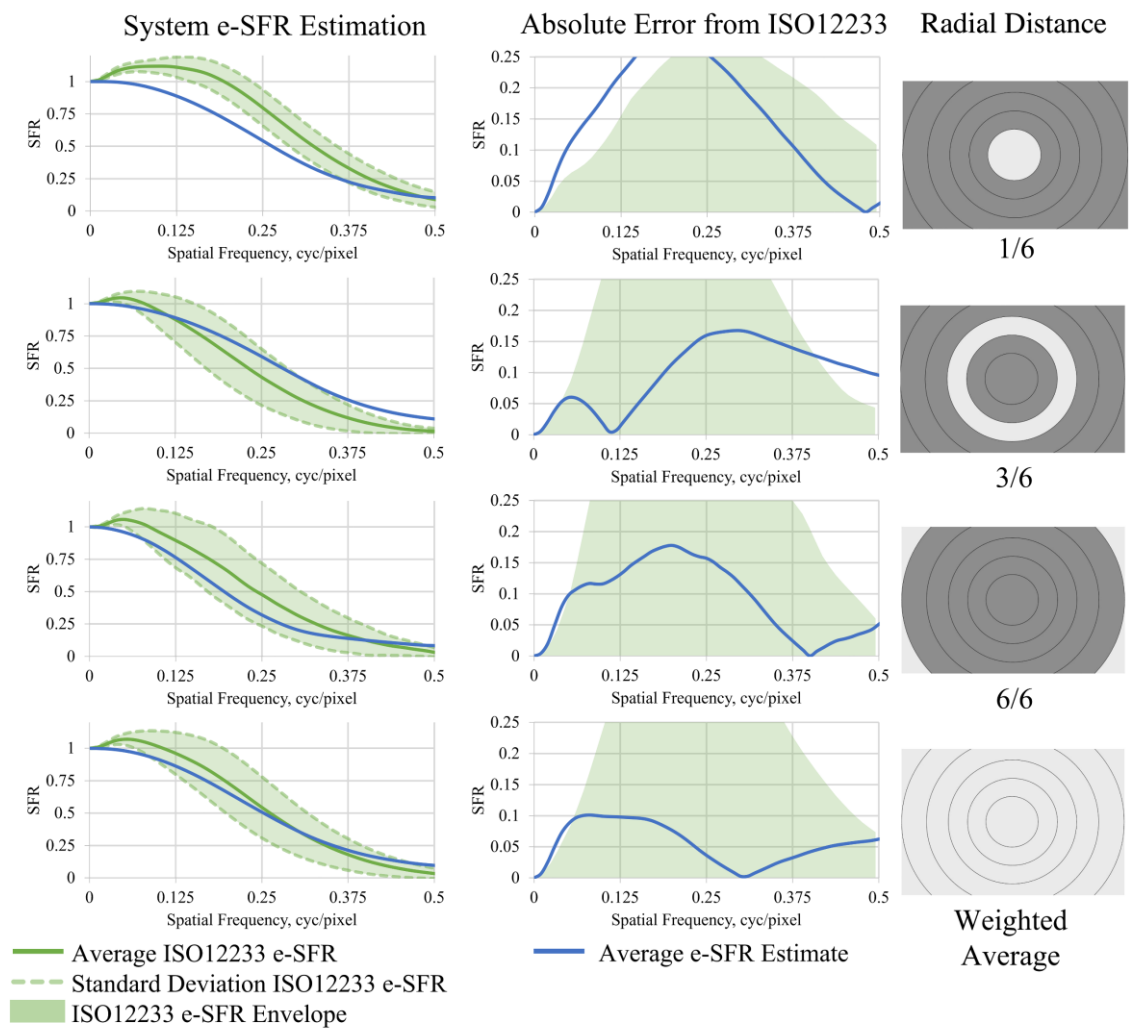


Figure 5.19 The Smartphone vertical system e-SFR estimation for three radial annuli out of six and a weighted mean of the entire field of view. The first column contains the estimated system e-SFR in relation to the ISO12233 e-SFR. The second column contains the absolute error between the estimated system e-SFR from the mean ISO12233 e-SFR. The third column contains a visual representation of the radial distance from which the data belongs [146].

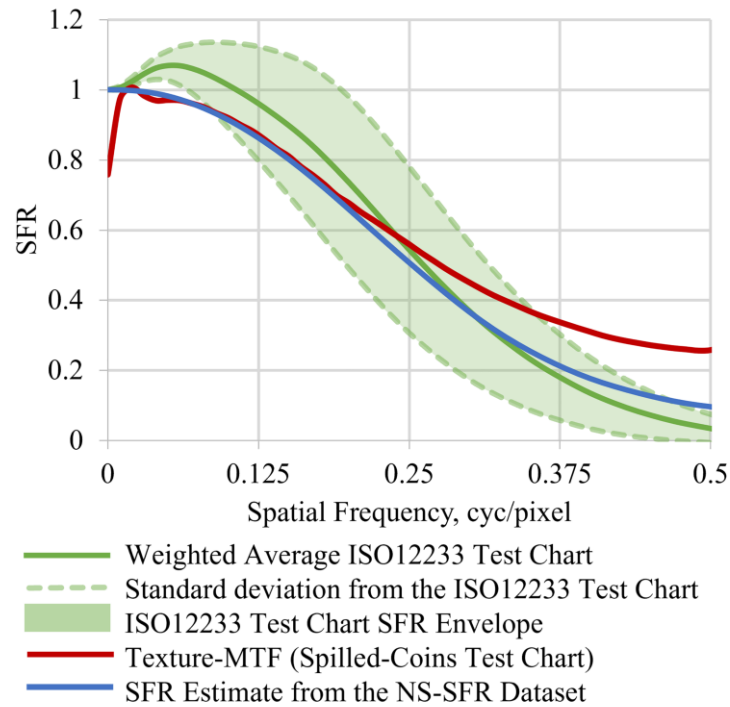


Figure 5.20 Weighted average vertical system e-SFR estimate from the entire Smartphone dataset. These estimates are compared to the ISO12233 slanted edge method and Spilled-Coins Texture-MTF [145, 146].

Scene dependency originating from non-linear processing is observed in the results. When isolating step-edges from test charts, adaptive processing (e.g., sharpening, denoising, compression) has a significant effect; chart edges are preserved and enhanced, resulting in a low-frequency boost in the e-SFR. Processing step-edges in complex natural scenes does not result in a sharpening boost. The inclusion of surrounding scene content and textures means that sharpening is not as effective on natural scene edges. Thus, the 25% Nyquist frequency metric has a greater absolute error than the higher frequency metrics, as seen in Figure 5.18. Denoising is also less effective in textured natural scene images than isolated test chart edges; hence, image noise further biases the estimation.

As with the DSLR 1 dataset (Figure 5.15), the smartphone camera dataset was segmented into three scene classes to investigate the scene dependency of system e-SFR estimates. Vertical e-SFRs for each dataset are plotted in Figure 5.21; horizontal system e-SFR estimations follow the same trends.

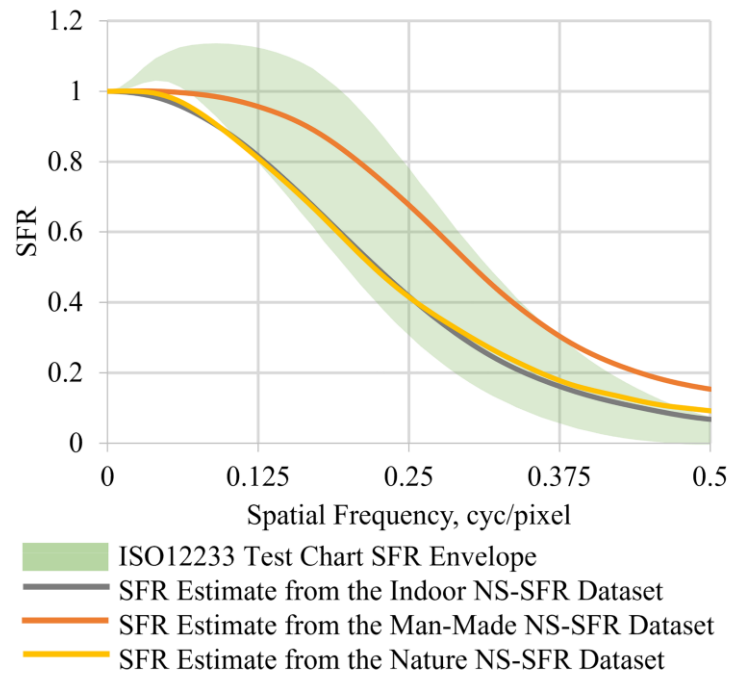


Figure 5.21 Weighted average vertical system e-SFR estimate from the Smartphone sub-datasets of indoor, man-made and nature scene classes. These estimates are compared to the ISO12233 slanted edge method standard deviation envelope [146].

Results show that individual class estimated system e-SFRs fall within ISO12233 e-SFR acceptable limits (i.e. the standard deviation envelope) but do not follow the boosted high-frequency trend of the ISO12233 e-SFR. The man-made scene class has a boosted mid-frequency, indicating stronger sharpening than the other two scene classes. Nature scenes do not have strong step-edges but have busy textures that do not respond well to sharpening. Indoor scenes may have plentiful step-edges, but, generally, they have lower and non-uniform illumination levels, higher gain settings, and consequently suffer more from image noise, prompting denoising. Denoising results in a slight blur in these indoor scenes, further reducing the system e-SFR estimate. High-frequency texture and high system noise prompt an ISP that results in a similar e-SFR system response.

5.4 Summary

This chapter used NS-SFRs (cf. § 3.1) derived from extensive image system datasets to estimate the system e-SFR. This estimation was achieved by selecting the most probable suitable step-edge inputs followed by further restricting the data to fall within the established edge and ROI parameters. Selection thresholds were set to minimise the NS-SFR variation without excluding a large amount of data.

NS-SFRs were grouped into six radial distance segments across the frame to evaluate optical performance variations across the field of view. This operation resulted in six system e-SFR estimations. The NS-SFR LSF FWHM distribution was thresholded per radial distance segment to select edges corresponding to the narrowest LSFs (i.e. the highest performance NS-SFRs). Hence, the NS-SFRs most likely measured from suitable step-edge inputs were selected, i.e. step-edges that contain modulation higher than the rendering capabilities of the system (cf. § 2.1.3 & 2.9). Three LSF FWHM distribution thresholds were tested, the 10th, 20th and 50th percentiles.

The resultant NS-SFRs were stored within a 4D coordinate system, which binned the data to produce a single NS-SFR per combination of coordinates. The data was further refined by isolating NS-SFRs measured from edge and ROI parameters that introduce low variation to system e-SFR. These parameter thresholds were determined in Chapter 4 (cf. § 4.4) and were verified in this chapter using NS-SFR data. The selected NS-SFR data was averaged in the spatial domain to derive successful system e-SFR estimates across the frame, and a weighted average of six local system e-SFRs gave a global estimate. The weights used were 1.00 for the central part of the frame, 0.75 for partway and 0.50 for the frame's corners. These zone weight values are the default values in the Imatest image evaluation software [74] but can be adjusted for specific applications.

Only a small yield of NS-SFRs were utilised for the system e-SFR estimation. The LSF FWHM distribution threshold first isolates approximately 10%/20%/50% of the edges. Secondly, parameter thresholds deselected a further 66% of data. When horizontal and vertical edge orientations were converted to sagittal and tangential, a further 71% of data was deselected.

Vertical and horizontal system e-SFR estimation was shown to be within the set limits of accuracy for near-linear systems. The frame weighted average, computed using the 10th

percentile, stays within one standard deviation of the equivalent ISO12233 e-SFR at 25%, 50%, and 75% of the Nyquist frequency. Sagittal and tangential estimates are not as accurate due to less data being available for the calculation. In this case, using wider NS-SFR LSF FWHMs increased data quantity, which in turn improved the estimate. That said, if the data quantity is available in the dataset, the 10th percentile of NS-SFR LSF FWHM distribution is recommended to obtain the most likely suitable step-edge inputs for a system e-SFR estimation.

The proposed method was most accurate when measuring the camera system with the telephoto lens. Telephoto lenses generally have characteristics that improve the estimation accuracy compared to wide-angle lenses, such as having little performance variation across the frame. In addition, wide-angle lenses generally have greater artifacts and distortions that negatively impact the estimate.

Results from the smartphone system indicate that highly non-linear systems produce a scene and processing-dependent e-SFR estimate. This was observed when dividing the smartphone system dataset into three scene categories. Further work is required to determine the extent of the relationship between system performance and non-linear signal processing.

Chapter 6 Evaluation of Image Dataset Size and Scene Classification for System e-SFR Estimation

Thus far, images with a variety of scene contents were used to estimate the system e-SFR. Such diverse inputs are appropriate in certain use cases, for instance, the comparison between consumer camera systems. Some situations, such as monitoring real-time camera performance for autonomous vision systems, may require input images specific to the application. In this chapter, the effects of restricting scene content as well as the number of images from which the system e-SFRs are estimated is investigated.

6.1 Dataset Size Evaluation Methodology

The rationale for choosing a particular system for this study is presented (cf. § 6.1.1), including system and file format considerations. The research is broken into three parts, firstly the scene classification (cf. § 6.1.2), secondly subdividing the dataset into subsets (cf. § 6.1.3), finally measuring the NS-SFRs and deriving system e-SFRs for each subset (cf. § 6.2). Figure 6.1 illustrates an overview of this study.

6.1.1 Camera System under Evaluation

The DSLR 1 RAW image dataset (cf. § 5.3.1) was divided into smaller subsets and specific scene categories for the purpose. The rationale for using the RAW sensor images rather than TIFF files was that the e-SFR estimates derived from RAW files, at 25% and 50% Nyquist frequency, have higher accuracy (cf. § 5.3.3). More importantly, the TIFF data showed signs of non-linearity in the demosaicing process (cf. § 4.3.2 & 5.2), which would potentially skew the results, since any variation should result from the quantity and *quality* of these isolated scene edge inputs, rather than scene and processing system behaviour.

The DSLR 1 system performance has considerable variation across the field of view. In contrast, DSLR 2 has a variation from a shallow depth of field, i.e. defocused edges. These optical performance variation characteristics are difficult to separate from the error from the number of images used in the estimate. However, the effect of a shallow depth of field on the e-SFR estimate becomes more problematic with fewer images, whilst the optical

performance across the field of view stays constant. Hence, DSLR 1 was used in this study. The estimations were colour coded based on the ISO12233 e-SFR standard deviation to account for the performance variation, highlighting the acceptable variation from an equivalent test chart measurement.

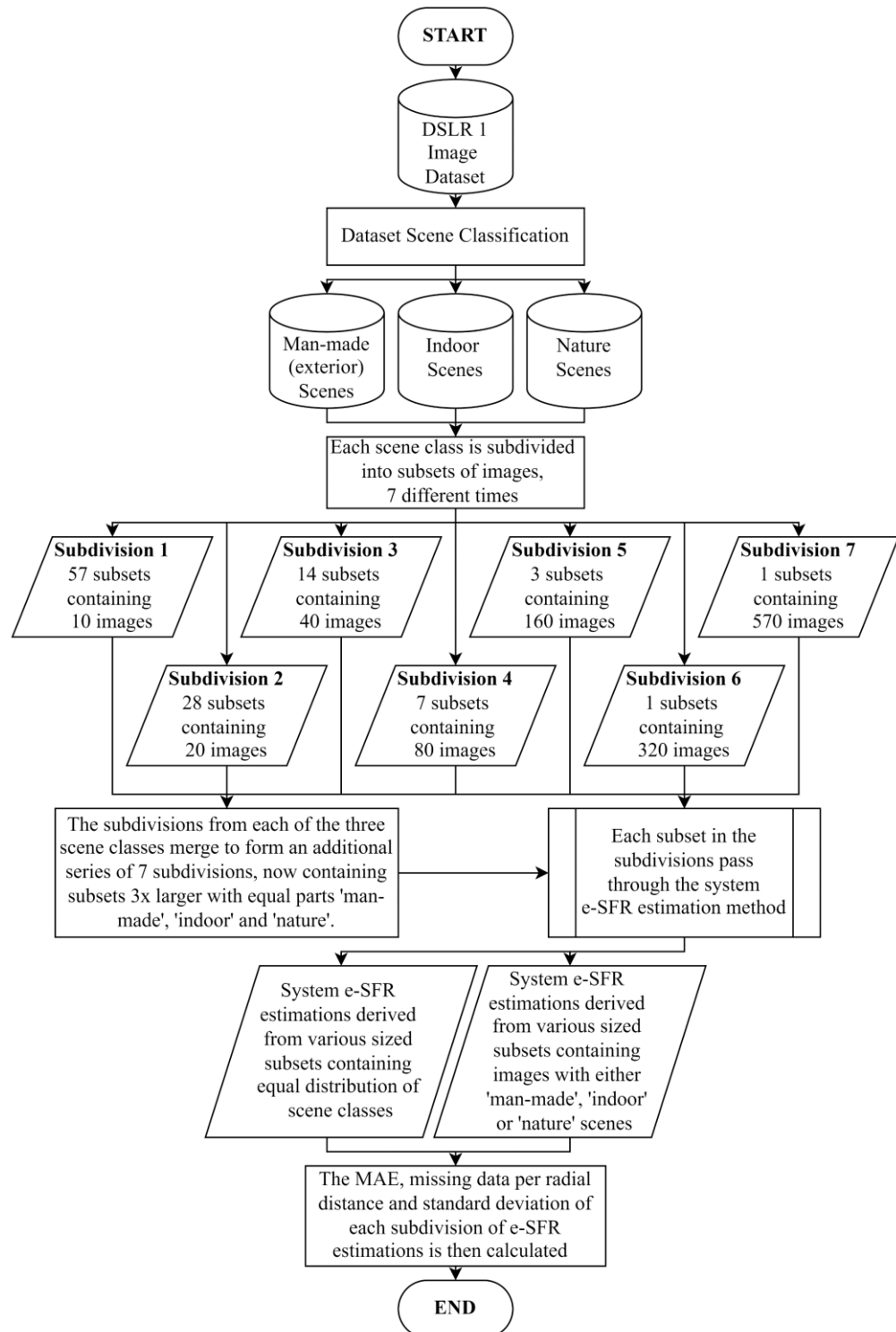


Figure 6.1 Methodology used to evaluate the accuracy in system e-SFR estimation when changing input image dataset size and scene content.

6.1.2 Scene Classification

Scene classification techniques are used in many areas of imaging science to distinguish scene types or find objects within images. The exact method for scene classification depends on the application, and in many studies, Neural Networks are trained using labelled training images to achieve this.

The AlexNet CNN [147] is a well-known and widely used network for object classification, accomplishing top-1 and top-5 error rates of 37.5% and 17.0%, respectively, which, at the time of publication, were considerably better than other state-of-the-art classifiers. AlexNet was trained using ImageNet [149], a database of 1.2 million training images, 50,000 validation images and 150,000 testing images, to provide five labels per image with over 1000 object categories.

Since AlexNet, there have been significant improvements in scene classification CNN architecture accuracy, with the introduction of GoogLeNet [150], VGG [151] and ResNet [152]. In addition, larger image databases have been shown to improve these object classifiers, such as implementing the 10 million image Places database [153].

Using a large number of categories was not a requirement. Instead, the aim was to classify scenes according to the general environment. Three categories were selected for this purpose:

- i) Man-made exteriors, urban or rural that include built structures.
- ii) Indoor scenes, including interiors of shelters, homes, and public buildings.
- iii) Nature scenes, which include landscapes, plants, and animals.

This simple classification allowed the breakdown of scenes in groups containing i) well-defined step-edges with low noise and low scene texture, optimal for the slanted edge method, ii) defined edges but with lower contrasts and lower illumination levels, thus containing higher image noise levels, and iii) less defined step-edges, containing more natural textures, suboptimal for the algorithm. Note that within each class, there is diversity in SNRs, scene objects and illumination.

These three categories are most relevant to the datasets captured for this project. More defined categories, such as people or roads, can be added for specific applications and NS-SFRs requirements.

Classifying images under these three scene categories was achieved through transfer learning [148] the AlexNet CNN [147]. Transfer learning is the process of fine-tuning a pre-trained neural network for a targeted task [148]. Rather than retraining all 25 layers of AlexNet, the last three layers were retrained, the *Fully Connected*, *Softmax* and *Classification Output* layers. This process keeps the foundation of the pre-trained network but trains it specifically for the task at hand. Resulting in less computation time than a full retrain of all 25 layers.

An extensive database containing 8,574 images was used in the transfer learning process. Each image was manually labelled under the desired categories. Half of these images were captured for this purpose, using a *Nikon D800* DSLR, a *Fujifilm X-T1* mirrorless and two smartphone cameras, the *iPhone 7* and *8*. The other half was taken from two open-access databases, SUN [154] and the Natural Scene Statistics in Vision Science [155, 156]. The SUN database is a collection of images taken from the internet to provide researchers with “annotated images covering a large variety of environmental scenes, places and objects”. The Natural Scene Statistics in Vision Science database is a collection of images photographed for research at the University of Texas at Austin. These images were captured using the *Nikon D700* DSLR with a *Sigma 50mm* lens.

Figure 6.2 describes the entirety of this database. The distribution of the diverse illumination sources is shown in the pie charts, and the specific scene locations are illustrated as bar charts. The ratio of the images categories indoor, man-made and nature was 38:20:27. The larger number of indoor images is to account for the greater diversity of interior scene contents.

Images were split 70:30 for training and validation purposes, respectively. The overall classification accuracy was high, at 96.35%. For a secondary validation, 100 DSLR 1 images per class were used to calculate the top-1 error rate. The top-1 error rate for the man-made classification was 18%, the indoor classification 6%, and the nature classification 2%.

Incorrectly labelled images were due to multiple factors, the most prominent source of error being scenes containing elements from multiple classes. Figure 6.3 displays examples of incorrect category labelling. The *certainty of prediction* is the validation output stated by AlexNet, i.e. the confidence that the image belongs to the predicted category.

These falsely labelled scenes may have caused some uncertainty in this study, particularly for the smaller subsets (cf. § 6.1.3) that may contain images that do not meet the intended

category, for example, Image e). In addition, uncertainty may have occurred when a scene contains two categories, for example, in Image b). The areas of nature structure have few isolatable step-edges, whilst most of the valuable step-edges in this example image come from man-made path structures. That said, the error was negligible mainly because of the low number of falsely labelled images. Also, although the subject matter of the photograph is another category, many falsely labelled scenes contain the edges and structures that belong to the classed scene type, for example, Images a), c) and d).

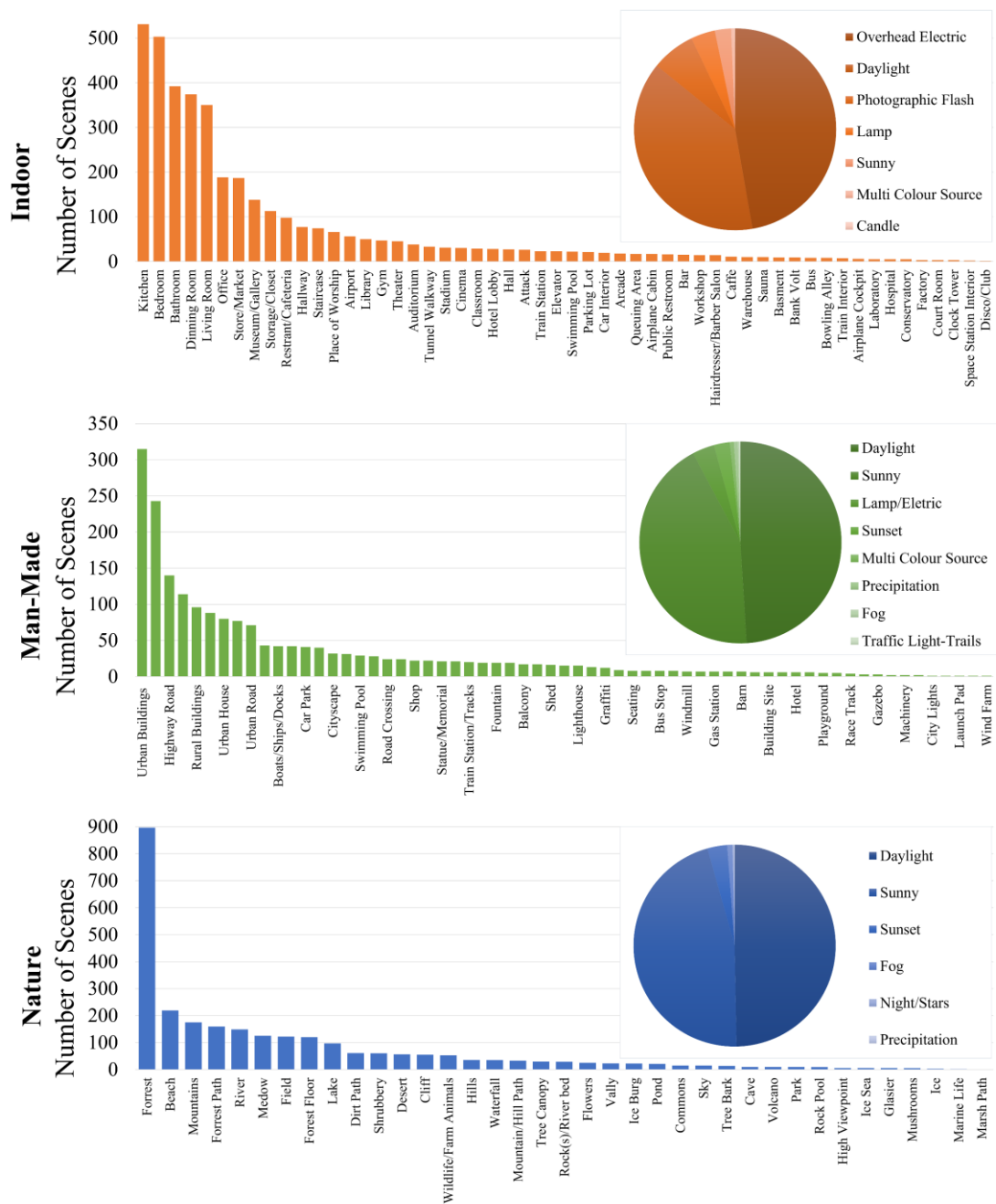


Figure 6.2 The distribution of the scene locations and lighting conditions of the indoor, man-made and nature datasets used in the AlexNet transfer learning process.






	Image	Category	Predicted Category	Stated Certainty of Prediction
a)		Indoor	Man-Made	97.33%
b)		Man-Made	Nature	99.95%
c)		Man-Made	Nature	99.84%
d)		Man-Made	Indoor	59.11%
e)		Nature	Man-Made	63.86%

Figure 6.3 This figure contains five examples of where the retrained AlexNet predicted the incorrect scene location category.

6.1.3 Subdividing the Dataset

The images in the three scene classes of the DSLR 1 dataset were divided into n groups containing m randomly selected images; these groups are referred to as subsets. This subdivision process was repeated seven different times, each time doubling m , as listed in Table 6.1, creating a logarithmic scale. As m increased, the number of subsets, n , decreased due to the finite size of the dataset. The images were randomly selected from the scene classes and were only used once per subdivision.

The subsets from the three scene classes were merged for each subdivision, forming an additional seven subdivisions. These merged subsets contained $3m$ images, equal parts man-made, indoor and nature. The number of subsets for each of the subdivisions remained the same. This division process is depicted in Figure 6.1.

These subsets were used to evaluate the number and type of scenes necessary for obtaining system e-SFR estimates comparable to the standard. The 10th percentile of the LSF FWHM distribution was used to derive the estimations. The absolute error at 50% Nyquist frequency was calculated between the estimate and the equivalent ISO12233 e-SFR for this evaluation.

DSLR 1 Dataset Subdivisions	Number of images per subset (m)	Number of subsets per scene class (n)
1	10	57
2	20	28
3	40	14
4	80	7
5	160	3
6	320	1
7	570 (all available)	1

Table 6.1 Seven subdivisions of the scene classed DSLR 1 dataset [146].

6.2 System e-SFR Estimation Results from Subdivided Dataset

The absolute error of the individual subset estimations was analysed with respect to subset size and scene class. Figure 6.4 provides a sample of this data analysis, showing the absolute error of the subdivided DSLR 1 RAW vertical system e-SFR estimates for radial segments 1/6, 3/6 and 6/6, and the weighted average of the entire frame. The data points are colour-coded to indicate whether the estimates are within, above, or below the ISO12233 e-SFR standard deviation limits. The MAE was calculated by averaging the subset error data points per subdivision. The MAE for each subdivision is plotted to show accuracy trends as the number of images processed through the proposed method increases.

Figures 6.5 and 6.6 plot the missing data and the standard deviation for each of the seven subdivisions against radial distance. These two figures provide the data for the mixed scene and separated scenes, respectively. Missing data is a percentage of subsets that did not provide an estimate for a given radial distance segment. The standard deviation plots show the spread of the estimates for each subdivision, providing a precision measurement. As subdivisions 6 and 7 only contain one subset, their standard deviations could not be calculated.

Three traits affect the accuracy of the estimation, which change across the field of view with the number of images used in the estimation:

- i) The quantity of edges available for the e-SFR estimate,
- ii) the scene dependency and suitability of the natural scene step-edges,
- iii) and the inherent performance variation of the system.

The centre of the frame, 1/6, has the least precise estimates, seen across all three scene classes, with high standard deviations. This is due to it containing few isolated ROIs, as previously shown, resulting in missing data. For subdivision 1, this missing data reaches up to 45% for the man-made and indoor classes and 70% for the nature class, i.e. it is difficult to estimate a centre radial distance system e-SFR using a small subset. When this segment contains suitable step-edges, a larger subset improves the accuracy and precision of the system e-SFR estimation, as seen in man-made and indoor scene classes. The e-SFRs estimated from the centre of the frame of the nature scene class, which has higher missing data and unsuitable step-edges, have no improvement in either accuracy or precision.

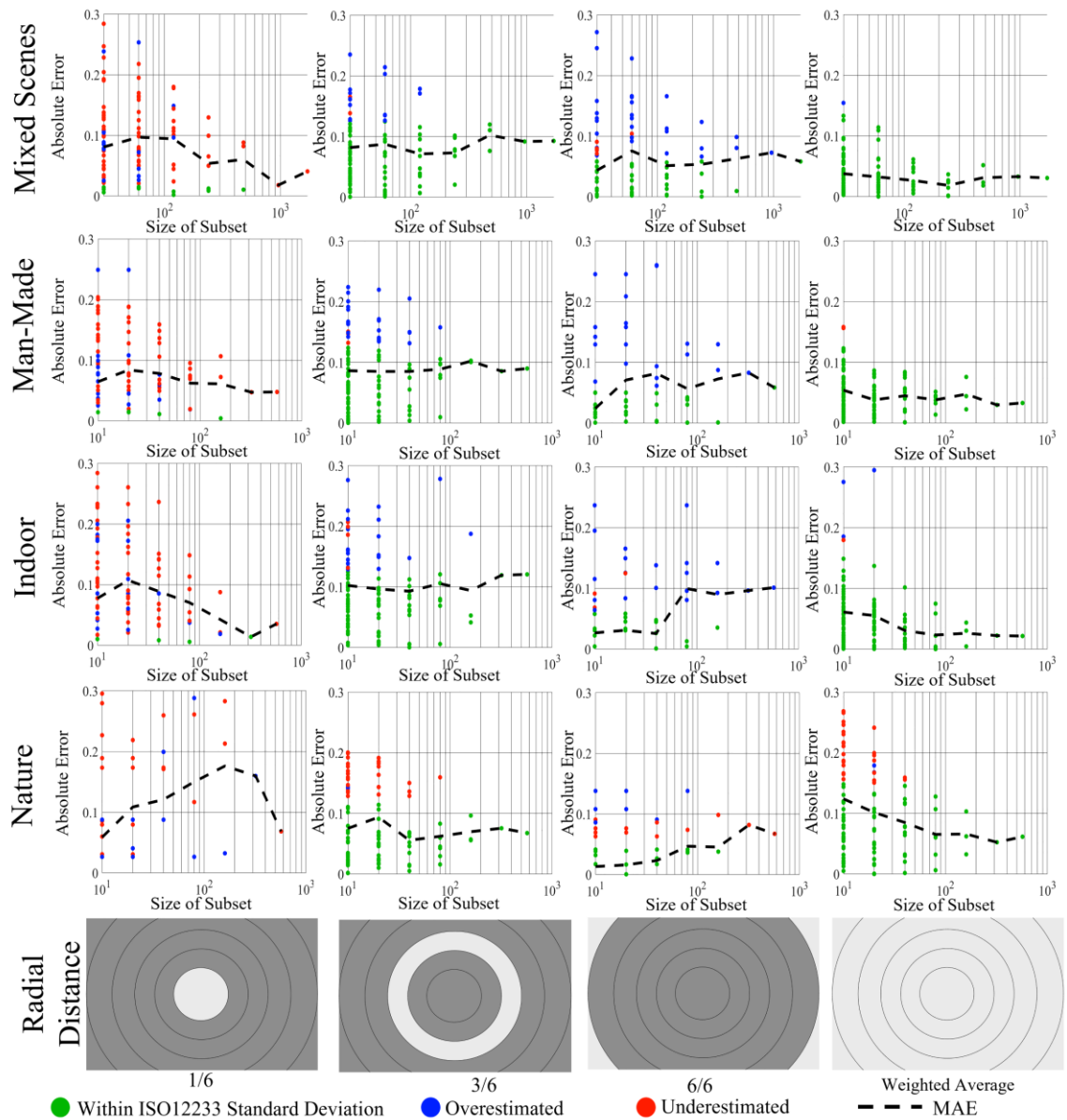


Figure 6.4 Plots depicting the precision and accuracy of the 50% Nyquist system e-SFR estimation, at radial distances 1/6, 3/6, 6/6 and the weighted average of the frame, using various sized subsets. This data is presented for the mixed scene subsets and the three scene classes, man-made (exterior), indoor and nature [146].

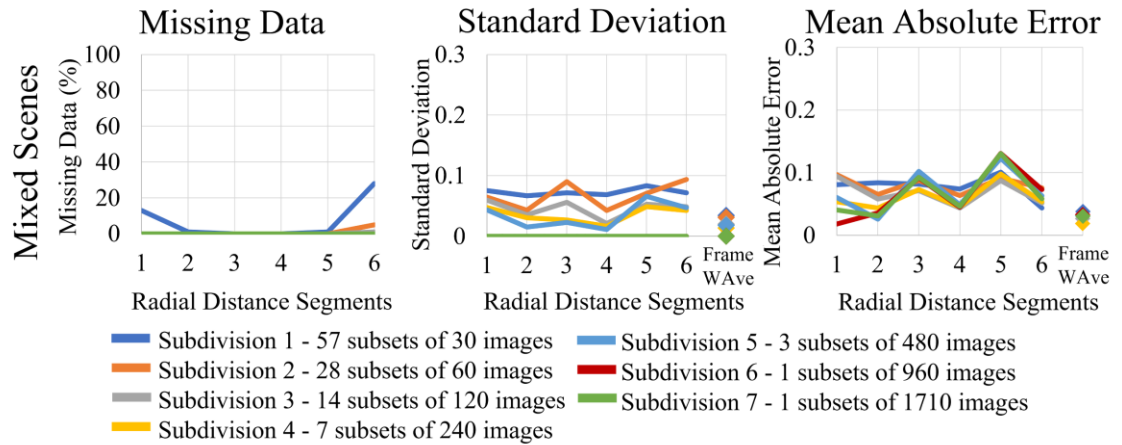


Figure 6.5 Plots illustrating the missing data and standard deviation across the seven DSLR 1 subdivisions against the radial distance segments for the mixed scene system e-SFR estimates.

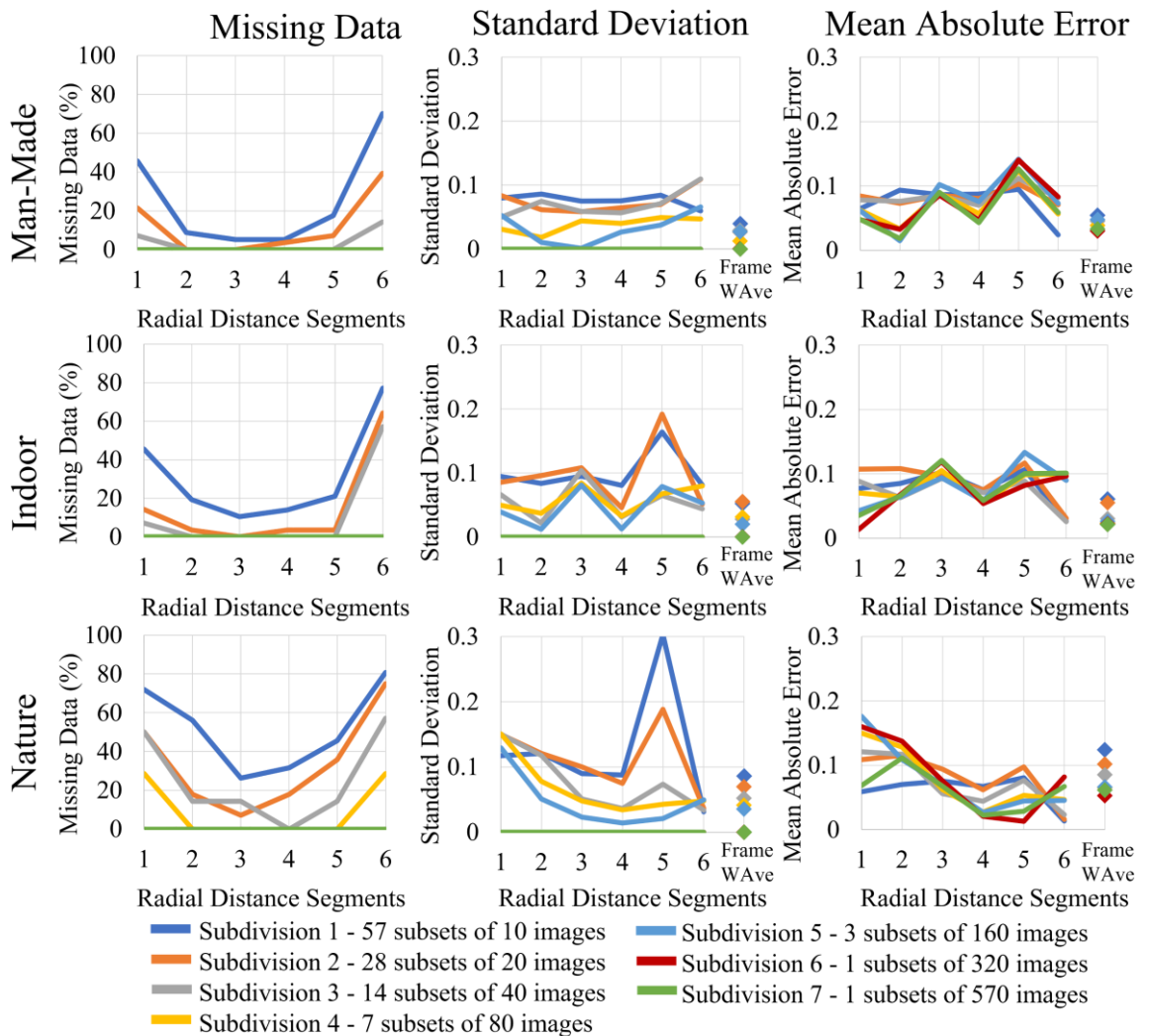


Figure 6.6 Plots illustrating the missing data and standard deviation across the seven DSLR 1 subdivisions against the radial distance segments per separated scene class [146].

This same observation is made for the corner frame segment, which has a maximum of 80% missing data due to few isolated ROIs as well as step-edges with lower suitability for all three scene classes. Thus, there is no improvement in precision with larger subsets. This is the lowest performance frame segment for a system due to high optical distortion and artifacts, such as chromatic aberrations, typically avoided in test chart based measurements. In contrast to the other frame segments, a larger subset decreases the accuracy, seen clearly in the nature class. More images provide more data to select suitable natural scene step edges, and the NS-SFR LSF FWHM distribution will impact the 10th percentile threshold. This is important in the corners of the frame, as artifacts such as chromatic aberrations can distort the measurement, increasing the distribution. Larger subsets have a higher likelihood to contain data with more artifacts, negatively impacting the estimate by increasing the distribution of the LSF FWHMs. In addition, where the corner edges are located impacts the estimation. The corners of the frame, particularly for a wide-angle lens, has the greatest performance variation. Small subsets may not be evenly distributed across each corner, resulting in subsampling the optical e-SFR range. Incorporating more images (thus, more edges) gives a more even distribution, providing an average of the range of e-SFRs, thus the increase in MAE for this radial segment. This is a challenging position in the field of view for the system e-SFR estimation, with missing data and problematic optical characteristics.

The partway region, 3/6, yields a high number of ROIs. As a result, fewer subsets contain missing data for this region. The standard deviation generally decreases with more images per subset, seen clearly in the man-made scene class. Although the standard deviation decreased, the MAE values stayed constant across the seven subdivisions for all three scene classes. In other words, when there are enough ROIs isolated, the average accuracy remains constant, but precision improves with more images. Also, note that the MAE stays within the ISO12233 e-SFR standard deviation, providing system e-SFR estimates comparable to the standard for all three scene classes.

The weighted mean system e-SFR estimation is a global measurement derived from plentiful data from every subset. Consequently, the standard deviation and MAE decrease with more images per subset, i.e. precision and accuracy improve. The accuracy improvement is minor in scenes with many suitable step-edges. However, the precision and accuracy improvements in suboptimal scenes with fewer suitable step-edges is far more significant.

Using 30 images of mixed scenes (the smallest subset) provided estimated system e-SFRs within one standard deviation of the ISO12233 e-SFR. Increasing the number of images does not significantly improve the MAE, but does increase the precision. That said, the centre of the frame is an exception, following the same central segment trends as previously discussed.

Providing a generalised conclusion of how many images are required for an accurate e-SFR estimate is impractical, as it is application dependent; instead, the number of edges should be determined using the following criteria:

- i) The scene content - Studying the scene classes, as expected, the system e-SFR is estimated with higher accuracy using small datasets containing images with well-defined step-edges. When using suboptimal scene inputs, it is beneficial to use large datasets and/or the weighted average estimate to assess system performance.
- ii) The e-SFR requirements - Suppose the application requires an e-SFR across the entire field of view. In that case, more images are required for accurate e-SFR estimates in the centre and corner frame segments. In contrast, if a single overall estimate is sufficient, fewer images provide an accurate e-SFR estimation across the tested scene classes.
- iii) The rate of change/variation of content between each image - In an application where the input images contain variety or are constantly changing, few images are required as the edges would be distributed across the entirety of the field of view. On the other hand, if the signal feed is static, increasing the number of images would not be beneficial.
- iv) The camera system - This study has used a high pixel resolution DSLR camera system. A lower pixel sensor and different optical characteristics would result in different step-edge yields and thus would require different numbers of images to achieve a robust e-SFR estimation.

Considering these criteria and assuming the ROI yields are similar to the DSLR 1 dataset, the following are examples of generalised practical applications. An automotive vehicle in an urban setting would have plentiful man-made structures. While driving, the movement of edges would provide edges distributed across the field of view. If an overall frame average is required, 10 images would be adequate to stay within the precision of the test chart e-SFR. This would provide three independent estimates every second, assuming 30 frames per second video feed.

For the same example system, but now requiring the e-SFR variation across the field of view, 80 images would balance precision and accuracy in the centre and corner frame segments. This would provide the e-SFR estimates every 2.67 seconds.

Another example would be a CCTV system. Such systems would have a static, or slow moving, video feed of man-made or indoor structures. Due to the inability to obtain evenly distributed edges across the field of view, a weighted average e-SFR of the frame would be most suited. Alternatively, the exact field of view location e-SFRs may be more appropriate than using azimuthal averages per radial distance segment.

6.3 Summary

The study presented in this chapter divided the RAW DSLR 1 dataset into smaller groups of images, classed by their scene location. These image subsets were then used as the input to the proposed method to estimate the system e-SFR. The aim was to assess the accuracy and precision of the estimation using small groups of similar images.

The AlexNet CNN [147] was used for scene classification. Transfer learning [148] was used to retrain the last three layers. This classifier grouped images according to the scene locations, either man-made, indoor or nature. These groups allowed the separation of scenes containing defined suitable step-edges that are optimal for the slanted edge algorithm, edges that are defined but less optimal and edges that are suboptimal. The accuracy of the retrained AlexNet CNN was shown to be 96.35%.

After the RAW DSLR 1 dataset was classified into the three categories, each class was subdivided seven times, creating a series of image subsets of seven different sizes. The subsets were grouped so that an additional seven subdivisions were created containing equal parts man-made, indoor and nature.

Estimating the system e-SFR from these subsets provided essential observations. The larger image subset increased accuracy and precision for radial distance segments found to have a low number of isolated step-edges, such as in the centre segment. In radial segments with suboptimal and fewer suitable step-edges, such as the frame corners, accuracy decreased with larger subsets due to optical distortion artifacts negatively impacting the step-edges. Partway radial distances were shown to contain a sufficient number of isolated step-edges; these areas maintained the same average accuracy across the subset sizes and for all scene

classes, but the precision increased with larger datasets. The weighted average was also shown to give the most stable result for small subsets, staying within the ISO12233 standard deviation limits, although accuracy and precision improved using larger subsets.

As expected, the data also showed that scene types with more suitable step-edges produce better system e-SFRs. Such scene classes allow estimates that maintain accuracy within one standard deviation of the ISO12233 using small datasets across all radial distance segments. The weighted average estimate provided the best results for suboptimal scene types, e.g., nature.

These results show that the number of images required in a dataset should be set based on the scene content, e-SFR requirements, the rate of change/variation of content between each image and the camera system.

Chapter 7 Discussion

This chapter contains discussions on the proposed methodology to measure NS-SFRs (cf. § Chapter 3) and estimate the system e-SFR (cf. § Chapter 4 & Chapter 5), the expected results from such an approach (cf. § Chapter 5 & Chapter 6), and potential applications. The strengths and caveats of this methodology and its results for the various camera systems tested are highlighted. This evaluation is made in the context of the broader subject area and background of system performance measurement (cf. § Chapter 2), identifying the novel research areas that were accomplished.

7.1 The Proposed Methodology for e-SFR Measurement

The overarching aim of the project has been to create an automated method that measures camera system performance directly from natural scene captures. The methodology presented in this thesis was developed across multiple studies and comprises two parts. The first presents a foundation that effectively extracts step-edges from natural scenes and utilises the ISO12233 e-SFR algorithm to produce the corresponding e-SFRs. These are referred to as natural scene derived SFRs (NS-SFRs). The second estimates the system e-SFR from this NS-SFR data, with the aim to provide a comparable measure to the laboratory-based ISO12233 e-SFR [5]. Although not all individual processes are original, the combination of processes and research behind developing the methodology is novel.

Edge detection plays an essential role in the framework, as it is the initial step of locating the edges within scenes. Two edge detection workflows were evaluated (cf. § 3.2). The first was a top-down approach, selecting many edges from the scene, then culling them to the required edge and ROI parameters for the e-SFR algorithm. The second was a bottom-up approach, directly selecting the appropriate edges. After reviewing both, the former was selected for the framework, which implements the Canny edge detector and subsequently applies processes to select and verify appropriate step-edges for the measurement. These additional selection stages become critical parts of the framework's success. They remove unsuitable edges and ROIs that introduce anomalies to the NS-SFR results. This approach, although robust, requires a high amount of further processing after the edges are selected. Future amendments to the edge detection would be required to make the framework more efficient. For example, a CNN trained using the data gathered from the Canny edge detection

workflow would result in faster step-edge extraction due to combining location, isolation, and verification processes, removing unnecessary computation to achieve the system e-SFR estimation.

Processes, such as the pixel stretching filter, developed to help isolate edges in ROI (cf. § 3.2.2) would remain in the CNN edge selection. Implementing the pixel stretching filter showed that, it not only improved the yield of edges isolated from a natural scene image, but also reduced the effects of image noise (cf. § 3.3.2.1) and ROI non-uniformity (cf. § 3.3.2.2). Drawbacks of this filter include streaking with noise or scene texture and amplifying the false-sharpening artifact found in some natural scene edges (cf. § 3.3.2.3). Also, applying the filter to tens of thousands of ROIs adds considerable computation time. A potential alternative approach to pixel stretching would be to compute only the ESF mask area and use this as a Hamming window in the ISO12233:2017 slanted-edge algorithm (cf. § 2.3.2), i.e. input only the masked area of the ROI. This would prevent the streaking and false-sharpening artifacts, but further studies are needed to implement such an approach and determine its impact.

It is important to note that the NS-SFR is not a camera system performance measure since it is derived from uncharacterised scene edges with unknown spatial frequency contents. Instead, it is a function of both camera performance and scene content, as observed in the NS-SFR scene envelopes (cf. § 3.6). Consequently, further processing is required to extract either the system e-SFR (cf. § 7.1.1) or develop a scene and processing dependent (SPD) performance measurement (cf. § 7.2.4).

7.1.1 Estimated System e-SFR Methodology

The principle of the proposed system e-SFR estimation is to select and group the NS-SFR data to reduce SFR measurement variation. This was achieved through thresholding the data to:

- a) group the expected optical performance variation across the imaging circle within radial distance segments,
- b) obtain the highest performance NS-SFRs per segment,
- c) select the NS-SFRs with input edge and ROI parameter ranges shown to introduce minimal e-SFR variation.

These selected edges from the NS-SFR framework cannot be guaranteed to be suitable for the slanted-edge method, i.e. edges that contain modulation higher than the capture capabilities of the system across the spatial frequency bandwidth (cf. § 2.1.3). Therefore the thresholds must allow for a stable system e-SFR estimation without excessively restricting NS-SFR data since suitable step-edges are rare in natural scenes.

The threshold to extract the highest performance NS-SFRs, and hence the most probable suitable step edges, was crucial for the proposed methodology. Three thresholds were tested (cf. § 5.1), being the 10th, 20th and 50th percentiles of the LSF FWHM distribution per radial distance segment. The 10th percentile has the advantage of selecting the narrowest LSFs, thus most likely to provide an e-SFR equivalent to the system performance. On the other hand, the 50th percentile has the advantage of more data inclusion. This is useful for obtaining estimates when the image dataset has few suitable step-edges, such as estimating e-SFR for sagittal and tangential orientations (cf. § 5.3.3).

For the most part, utilising a percentile threshold to obtain the highest performance data works as intended. However, it should be noted that it is not effective if all input edges are suitable, i.e. all the measurements provide an e-SFR, as taken from test charts. In such cases, the threshold will isolate the highest optical performance measurements, resulting in an overestimated e-SFR. The probability of this becoming a real-world problem is low when utilising natural scenes, as shadows, rounded corners and depth of field will incorporate unsuitable step-edges that are subsequently filtered out.

Variation is further reduced by selecting the NS-SFR data measured from the edge and ROI parameter ranges shown to be stable (cf. § 4.4); on average, this decreases the amount of data by 66% (cf. § 5.3.2). As a result, this approach produces a low data yield for the system e-SFR estimation; only 3.41% of the total isolated ROIs are utilised. Nevertheless, the selected data is robust for the estimation process.

The proposed methodology was applied to estimate the system e-SFR of two DSLR systems (DSLR 1 and 2), both using the *Nikon D800* body, but each with a different lens system. The first lens was a *Nikon AF-S 24-70mm f/2.8G IF-ED* at 24mm and the second a *Carl Zeiss 135mm f/2 Apo-Sonnar T**, with both lens apertures set at $f/4$. They were chosen because they have significantly different optical characteristics (cf. § 5.3.1).

To determine e-SFR estimation accuracy, limits were set according to \pm one standard deviation of the test chart measurements, per radial distance segment, and for the entire frame. These limits reflect the traits of each optical system. The wide-angle lens has relatively broad limit boundaries due to the considerable variation in optical performance across the imaging circle. In contrast, the telephoto lens has narrow limits due to the lower variation in optical performance. The range of these standard deviation limits can impact the perceived accuracy of the estimates. For instance, the estimate for the DSLR 1 system may be within one standard deviation of the equivalent test chart measurement, and the estimate for the DSLR 2 system may be outside the accuracy limits. But the MAE between the DSLR 2 estimate and the equivalent test chart measurement may be lower than that of DSLR 1. Therefore, to assess accuracy, both the standard deviation of ISO12233 e-SFRs and the MAE from the average ISO12233 e-SFR per radial distance were considered.

The estimated system e-SFRs from both systems were, for the most part, within one ISO12233 standard deviation, although a number of them were overestimated. The DSLR 1 results, calculated using the 10th percentile LSF FWHM, show higher accuracy in radial segments with narrow envelope variation than segments with broad envelopes, such as the centre of the frame, staying within a delta of 0.05 from the mean ISO12233 e-SFR at 50% Nyquist. The partway region, which naturally has a broader performance variation, gave an absolute error of approximately 0.10 to 0.15 (depending on whether the mosaiced or demosaiced image was used). This observation is reflected throughout the DSLR 2 system e-SFR estimations using the 10th percentile threshold. The tighter performance tolerances of the telephoto lens provide delta values under 0.06 measured at the 50% Nyquist frequency across all radial segments. Thus, the proposed methodology works better for systems, or radial segments, with little optical performance variation. The weighted average of the local system e-SFR estimates across the frame of these near-linear camera systems are within one standard deviation of the equivalent ISO12233 e-SFR measurement (cf. § 5.3.3).

Expanding the LSF FWHM distribution threshold to the 20th or 50th percentile can reduce the delta value in radial segments with wide optical performance variation, providing an estimate closer to the mean ISO12233 e-SFR. However, this would cause underestimation when using an optical system with narrow performance tolerance. This disparity relates to the NS-SFR measurements being affected by system performance and scene contents. Camera systems with little optical variation will result in a high threshold when using the 10th percentile of the LSF FWHM. As a result, there would be a minimal overlap between

NS-SFRs containing only the system performance and those also containing scene content. In contrast, wide variations in performance will result in a significant overlap, making the threshold position more problematic. Again, the 10th percentile of the LSF FWHM would ensure that the system performance is being selected; however, it may also only select the higher performance range of the optical variation. Using a lower threshold (20th or 50th percentile) may incorporate the expected optical variance but also adds uncertainty to the measurement. It is unknown if the performance variation in the estimation is due to the optical system or the inclusion of scene contents. Thus, the LSF FWHM distribution threshold must be carefully considered when implementing, as it must also correspond to the optical characteristics of the system. This work concluded that the 10th percentile was the most appropriate threshold in most applications for excluding the scene content component of the NS-SFR.

When implementing this methodology to highly non-linear systems (cf. § 5.3.4), the resulting estimates indicated scene dependency. The weighted average estimation for the Smartphone dataset has a low-frequency curve comparable to the texture-MTF results, without the typical sharpening lobe. The low-frequencies show bias introduced by image noise. This e-SFR estimation does not resemble the edge test chart e-SFR, suggesting that sharpening and denoising are applied on the test chart, whilst pictorial natural scene content is not processed as strongly. The results further indicate scene dependency after segmenting the dataset according to scene location. Additional studies are required to develop a scene and processing dependent SFR (cf. § 7.2.4) from the NS-SFR data.

7.1.2 Methodology Improvements

Although the method is novel and robust for measuring near-linear camera systems, several areas could be improved.

The focus distance is a known factor that can change the optical performance, but it is not taken into account in the method. The focus distance is not an easily obtained measurement, it is not commonly present in an image file's EXIF data, and when it is, it is rarely accurate. The concept of obtaining the lens's focus distance by utilising a single image depth map prediction [111–115] was reviewed in Chapter 2 (cf. § 2.9). The focus distance could be estimated by pairing a depth map and the location of the highest performance step-edges. This was not pursued in this research because single image depth map predictions are

currently not accurate enough for such purposes. They do not provide quantifiable distances or relative distances across multiple images. Further work in this field is required before this method can be implemented into this workflow.

Motion blur is another issue that has not been addressed in the current methodology. Outside a laboratory-based measurement, natural scenes captured using a slow shutter speed can cause motion blur; this introduces a negative error in the estimated system e-SFR. One easy solution would be to remove images captured with low shutter speeds, but this would also remove images shot on a tripod. That said, using the highest performance NS-SFRs from multiple images mitigated this issue. Motion blur is a topic to be studied for applications that use single images and are prone to it. A possible solution would be to utilise motion blur detection and correction (cf. § 2.9) [116] for applications where motion blur is in a predictable direction at a known speed (such as autonomous vehicles (cf. § 7.4.3)). Further research into how, and if, motion blur can be factored into the estimation would also be beneficial.

Illumination non-uniformity is a problem in the slanted-edge method (cf. § 2.5.2), but as discussed, the pixel stretching filter helps reduce its effects on the e-SFR. In this work, only low gradient non-uniformity were considered; more complex non-uniformities need to be studied, for instance, uneven shadows on one side of the ROI. An alternative solution would be to use an approach similar to the non-uniformity correction in the Imatest software [77]. Imatest's method estimates a first-order fit on the light side of the ROI, as the light side has a higher SNR, then divides the edge with this fit. This successfully removes low gradients caused by incorrectly lit test charts. Research into using higher-order fitting functions may be valuable for more complex non-uniformities found in natural scenes.

Rather than selecting the NS-SFR data that most likely provide a reliable performance measurement, an alternative approach would be to estimate the e-SFR using statistical analysis of the distribution of NS-SFR data and the NS-SFR maxima. A large image dataset (2000+ images) would provide many captured scene edges for statistical analysis. In some respects, this concept can be more robust than the approach taken in this project because there are no predetermined judgements/thresholds. That said, it would also be more computationally expensive. It would require knowledge of multiple systems, the ISP and scene content to determine a statistical approach that would produce an estimation representative of the ISO12233 e-SFR.

7.2 Implementation

7.2.1 Code Optimisation to Improve Usability

All code has been written using MATLAB, utilising parallel computing to utilise all the cores of the available Central Processing Unit (CPU); this code is open access and is available on a *GitHub* repository (cf. § Appendix E). Further work is required to improve and fine-tune the usability of this code and the implementation of the proposed method.

The method is coded as presented in this thesis, first measuring the NS-SFRs from a dataset and then estimating the system e-SFR. For any practical implementation, the code must be optimised. The system e-SFR estimation is a highly selective process that yields small numbers of suitable natural scene step-edges. As a result, assuming that the NS-SFR data is not required for the application, such as in SPD measurements (cf. § 7.2.4), and a system e-SFR is the objective, this procedure is inefficient with 96.59% of the isolated data being deselected. The current computation times are, on average, 20 minutes per image (based on 36-megapixel image files and single CPU core computing).

Thus, the first step in implementing the proposed methodology for real-world applications would be to restructure the code to select the utilised edges used in the estimation (3.41% of the currently selected data). However, such an efficient yield of utilised edges would not be possible with the current approach, as the full LSF FWHM distribution is required to select the highest performing NS-SFRs. Instead, the data with the correct edge and ROI parameter values could be isolated at the beginning of the process, reducing the amount of unused NS-SFR processing by approximately 66%. After implementing this simple optimisation, the computation time is expected to reduce to a few minutes per image, further improved by applying additional hardware acceleration, specifically Graphics Processing Unit (GPU) acceleration.

Additional optimisation could be achieved by isolating the top 10%/20%/50% sharpest edges from each image or dataset prior to e-SFR estimation, removing the need for the LSF FWHM distribution threshold. This could be carried out in two ways: the first would use a gradient-based edge detection algorithm to locate the edges of required sharpness. The Canny edge detector could be used for this purpose. Instead of detecting an extensive range of edge gradients, appropriate thresholds could be applied to find the highest gradients. However,

work is required to determine thresholds that output the 10%/20%/50% of the sharpest edges. Secondly, a CNN could be employed and trained to select the required edge gradients and profile, significantly improving the computation times. The problem with both methods above would be their system dependent nature; the sharpest edges would be system dependent. Further work is required to assess such approaches and obtain an optimised algorithm, which is also flexible and adaptable for multiple systems.

7.2.2 Recommended Settings and Thresholds

All thresholds in the proposed methodology, both in the NS-SFR framework (cf. § Chapter 3) and the system-SFR estimation (cf. § Chapter 5), must be set appropriately for the dataset, system, and system settings for successful implementation. The application for the method also is important, as the desired output may influence the settings and thresholds.

The file format of the image dataset from which the e-SFR is estimated directly impacts the estimate. Demosaiced image files are suitable for the estimation, provided the ISP is linear, or close to linear. Alternatively, the RAW (mosaiced sensor) image can be used to ensure linearity. Compressed image files, such as the JPEG files, and images from non-linear pipelines, are not recommended without further study (cf. § 5.3.4).

Studies conducted in this thesis (cf. § 6.2) have shown that the size of the input dataset and the captured scene contents combine to determine the accuracy and precision of the estimated e-SFRs. If the application allows, an image dataset containing 1000 or more images is recommended, as it was shown to increase the estimate's accuracy. When such large datasets are not available, the accuracy becomes dependent on the type of scenes contained in the dataset. Fewer images are required in the dataset to provide accurate estimates across the frame when the image contains well defined, suitable step-edges. In contrast, a weighted average would be recommended for small datasets (10-80 images) containing scenes with suboptimal and few suitable step-edges.

Crucial thresholds in the NS-SFR framework that should be optimised for each system include the thresholds set for proximity filter (cf. § 3.3.2) and noise floor (cf. § 3.3.2 & 3.3.3). The threshold for the proximity filter is currently set to deselect neighbouring edges within a 5 pixel distance. This is appropriate for the *Nikon D800* DSLR, as the modelled MTF corresponds to an ESF distribution of no more than 4 pixels (cf. § 3.3.1). This threshold can be decreased or increased depending on the modelled resolving power of the test system

(cf. § 2.2.3). The noise floor should be set to best suit the camera and ISO gain settings of the images in the dataset used for the estimation, currently set at the fixed threshold of a 0.02 normalised pixel value change. Ideally, this value should be adaptive on an image-to-image basis.

In the e-SFR estimation methodology, there are thresholds optimised for the system under test and the application. For example, the previously discussed LSF FWHM threshold should be set to best suit the dataset and optical characteristics of the system. Other factors that should be considered are the weights used in the frame average and the number of radial segments required for the application.

When the e-SFR variation across the field of view is not required, fewer images in the dataset are necessary to provide estimation accuracy comparable to the ISO12233 system e-SFR. In contrast, if a measurement of performance variation across the field of view is required, larger image datasets are needed to obtain accuracy across all radial segments; more radial segments mean more edges are necessary, preferably evenly distributed across the frame. Alternatively, when the performance variation is required, but the dataset is small or contains captured scenes with suboptimal and few suitable step-edges, the number of radial distance segments should be reduced, such as dividing the frame into three segments rather than six. Decreasing the number of radial distance segments is inappropriate for applications involving sagittal and tangential e-SFRs, as they are commonly plotted using a performance metric against the radial distance (cf. § 2.6.2). Furthermore, the appropriate sagittal and tangential edge orientations are not common in captured natural scenes (cf. § 5.3.3). Hence, either a large dataset is required, or as mentioned above (cf. § 7.1.1), if such a dataset is not available, the LSF FWHM distribution threshold should be set to incorporate more edges, using 20th or 50th percentiles.

7.2.3 Improvements to the ISO 12233 Algorithm

The goal of the proposed method is to estimate the ISO12233:2017 e-SFR from natural scene captures. The NS-SFRs were calculated using the standard slanted-edge algorithm. The only alteration was fitting a 3rd order polynomial to the edge slope projection (Equation 2.29) to accommodate for curved edges. This higher-order fitting was essential to reduce bias in the measurement [86, 135], since 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 has been shown to improve the e-SFR measure in multiple papers [61, 64, 84, 157] (cf. § 2.7.1). These publications provide results with reduced bias introduced by image noise, edge angle variations and ROI size variations. The approaches in 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 e-SFR measure [61]. In contrast, the Reverse Projection method remodels the ESF resampling process, giving more stable results, across large edge parameter ranges, than other slanted edge method algorithms [64]. However, the resultant Reverse Projection e-SFR is consistently underestimated compared to ISO12233 algorithms, including *sfrmat2* and the Imatest implementations. Applying such alternative methods would significantly improve the NS-SFR stability, allowing more ROIs to estimate the system e-SFR. However, if employed, the accuracy and impact on the standard should be understood.

A revision of the ISO12233 standard is scheduled for release later this year (in 2022); this will be the fourth edition of this standard. The additions and changes that are intended in this revision will include [78]:

- Advanced edge fitting, using a 5th order polynomial fitting function.
- SFR correction to account for edge angle variation, in the same manner as Imatest [61] (cf. § 2.7.1).
- Improvement to the signal-processing window, replacing the Hamming window with a Tukey window.
- Non-uniform illumination correction, as Imatest includes with their software [77] (cf. § 2.5.2).
- The inclusion of an e-SFR acutance measurement.

Once released, these advances in the ISO12233:2022 [78] would benefit future revisions of the NS-SFR and system e-SFR estimation methodology. Research into the benefits of the newly proposed edge angle variation and non-uniformity corrections on natural scene ROIs would be of particular interest.

7.2.4 Scene Dependant e-SFRs

Most modern cameras incorporate ISPs, which can be treated as a black box. Results from a smartphone system incorporating highly non-linear ISP indicated that the e-SFR estimate is SPD. Further work is required to establish correlations between specific scene contents (scene edge types, texture, and noise), the NS-SFRs and the resulting e-SFR estimates. Such an approach moves away from linear system theory, as assumed in traditionally laboratory e-SFR measurements, potentially considering the adaptive nature of systems in future performance measurement solutions.

Scene content can be assessed using spatial based Natural Scene Statistics (NSS) on a global or local image level. Simple examples include the mean, median and skewness that can be used to assess luminance. Entropy can quantify image structure, while the root mean square (RMS) contrast can be measured using the standard deviation and mean pixel value of the image or ROI [50, p. 388].

Natural scene metrics can also be used to measure the spatial content of captured images, for example, the busyness metric [158], a texture feature metric [159], noise metric [160], and blur quality score [161]. Established aesthetic analysis can also be used or adapted for this purpose. The metric established by Aydin et al. [162] could help determine scene and performance dependency:

- Sharpness metric – measurement of absolute contrast magnitude of the in-focus region of the image.
- Clarity metric – quantifies if the image has an identifiable subject region.
- Depth metric – an index of the dominant spatial frequency at the non-sharp image regions.
- Area of in and out of focus regions.

NSS and other metrics should strictly reflect the factors affecting the system performance and the slanted-edge method. Sharpness, noise and the focus point are important metrics. Also, the busyness metric [158], or a texture feature metric based on the Hough transform [159], should be considered, as the level of texture directly correlates with the number of step-edges extracted. In addition to global measurements, a study that breaks down localised scene content combined with the localised NS-SFRs would further this work.

Research into the e-SFR estimate from different edge contrast ranges, as well as using grouped LSF FWHM distribution thresholds (e.g., percentile ranges 10th to 20th, 20th to 30th, 30th to 40th etc.), would also be beneficial. Such studies should employ datasets containing variations in natural scene content, even distributions of edge contrasts, the RAW and processed file formats, allowing linear and non-linear signal processing behaviour to be investigated.

Such studies would provide insight into the nature of scene-and-process dependent camera ISP, determining whether e-SFRs estimated from natural scenes would be a more suitable non-linear system measurement than the traditional laboratory-based ISO12233 e-SFR.

7.3 Comparison to the PSF CNN Approach

As reviewed (cf. § 2.8.4.1), a natural scene derived MTF was proposed by Bauer et al. [15]. To summarise, their approach used a neural network to estimate the PSF across the frame directly from natural scenes, resulting in the sagittal and tangential MTFs that were calculated with high accuracy. This methodology is referred to as the PSF-CNN approach, whilst the proposed methodology from this thesis is referred to as the slanted-edge approach.

In this section, the reported accuracy of the PSF-CNN approach was compared with the accuracy of the proposed slanted-edge approach. Ideally, the accuracy of the two methodologies should be assessed using the same datasets, across multiple systems, for a fair in-depth comparison. Instead, the aim here was to provide a context of the accuracy between alternative approaches for pictorial natural scene derived MTF/SFR.

The paper presenting the PSF-CNN approach published the results and the absolute accuracy for ten DSLR systems, each using the *Canon 5DS R* (containing a 50.6 megapixel sensor, with a pixel pitch of 4.13 μ m) with different lenses. Their study included a *Canon EF 24mm f/1.4L USM*, a *Canon EF 135mm f/2L USM*, and a *Carl Zeiss 100mm f/2 Makro-Planar T**. Although not identical lenses to the DSLR 1 and DSLR 2 systems (cf. § 5.3.1), they provided similar focal length characteristics to compare. DSLR 1 used a *Nikon 24mm* lens at $f/4$, comparable to the wide-angle Canon 24mm. The DSLR 2 used a *Zeiss 135mm* lens set at $f/4$ that was compared to the telephoto characteristics of the *Canon 135mm* and the optical glass anti-reflective T^* coating qualities from the *Zeiss 100mm*. The closest aperture setting used

in their testing was $f/5.6$. This aperture discrepancy is negligible in this comparison, as it is a test of accuracy with their equivalent test chart measurements.

Bauer et al. used the 10, 20, 30 and 40 line-pairs per millimetre (lp/mm) to assess their PSF-CNN approach. Thus, the DSLR 1 and DSLR 2 results were compared using these same metrics. Their results were given in graphical form; therefore, the data was approximated from the provided graphs. Also, as the azimuthal average was used in both approaches, only this averaged data was compared. All tested DSLR sensor sizes are identical, making radial distance comparison straightforward; the data was compared at radial distances 3.6mm, 10.8mm and 19 mm (equivalent to 1/6, 3/6 and 6/6 radial segments).

The slanted-edge approach is best suited to estimating the e-SFR from horizontal and vertical edges (cf. § 5.3.3). This is because there are more of these orientations available from natural scenes than sagittal and tangential edges. That said, to provide a direct comparison to the PSF-CNN approach sagittal results, the sagittal system e-SFR estimate was also used alongside the vertical e-SFR estimations.

The input images from both approaches were the green channel of the mosaiced RAW files, except for the DSLR 2 dataset that used the TIFF file format.

Figure 7.1 depicts the peak absolute error across all radial distances for the four spatial frequencies, 10 to 40 lp/mm. In Figure 7.2, the absolute error for spatial frequencies 10 to 40 lp/mm at the three radial distances for the slanted-edge approach (vertical and sagittal orientations) and the PSF-CNN approach are plotted. The graphs are grouped to compare the approaches using similar systems directly.

This comparison has shown that the proposed vertical e-SFR slanted-edge approach is level with or, in some cases, better than the alternative PSF-CNN approach. The peak vertical e-SFR estimations are approximately half of the comparable PSF-CNN approach estimates. There are weaknesses to the slanted-edge approach, for example, converting the wide-angle lens system orientation to obtain sagittal e-SFRs caused peak absolute errors to be 26% higher than that of the PSF-CNN approach.

Both methods are shown to provide estimates with higher accuracy when implemented with high performance systems. The telephoto lenses, for example, show a lower measured error than the wide-angle lenses. For the telephoto lens systems, the peak error of the vertical

slanted-edge estimate is half of the equivalent PSF result, whilst the sagittal e-SFR estimate peak error is equal to that of the PSF method.

In the publication presenting the PSF-CNN approach, other high-performance lenses were studied, for example, the *Zeiss Otus 55mm f/1.4 APO-Distagon*, which had an absolute error peak of 0.05. It must be noted that this *Zeiss* lens (and the *Canon 24mm*) were used in the PSF neural network training process [15], thus may bias the accuracy of the result for these lenses.

This is not a comprehensive comparison between the two methods. The number of images and the scene types used in the analysis of the PSF-CNN approach are not available from its publication. The slanted-edge approach used extensive datasets of 1866 images to achieve these estimates. Further work is required to provide a fair study of these different approaches. Such a study should use an independent test system.

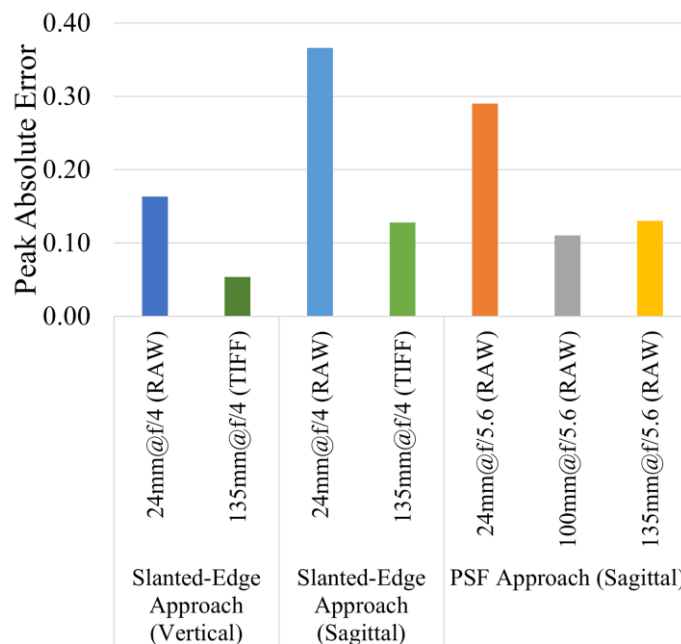


Figure 7.1 The peak absolute error from the slanted-edge (as presented in this thesis) and PSF-CNN (as published by Bauer et al. [15]) approaches, at 10, 20, 30 and 40 lp/mm across the entire frame. This comparison is between similar, but not identical, systems.

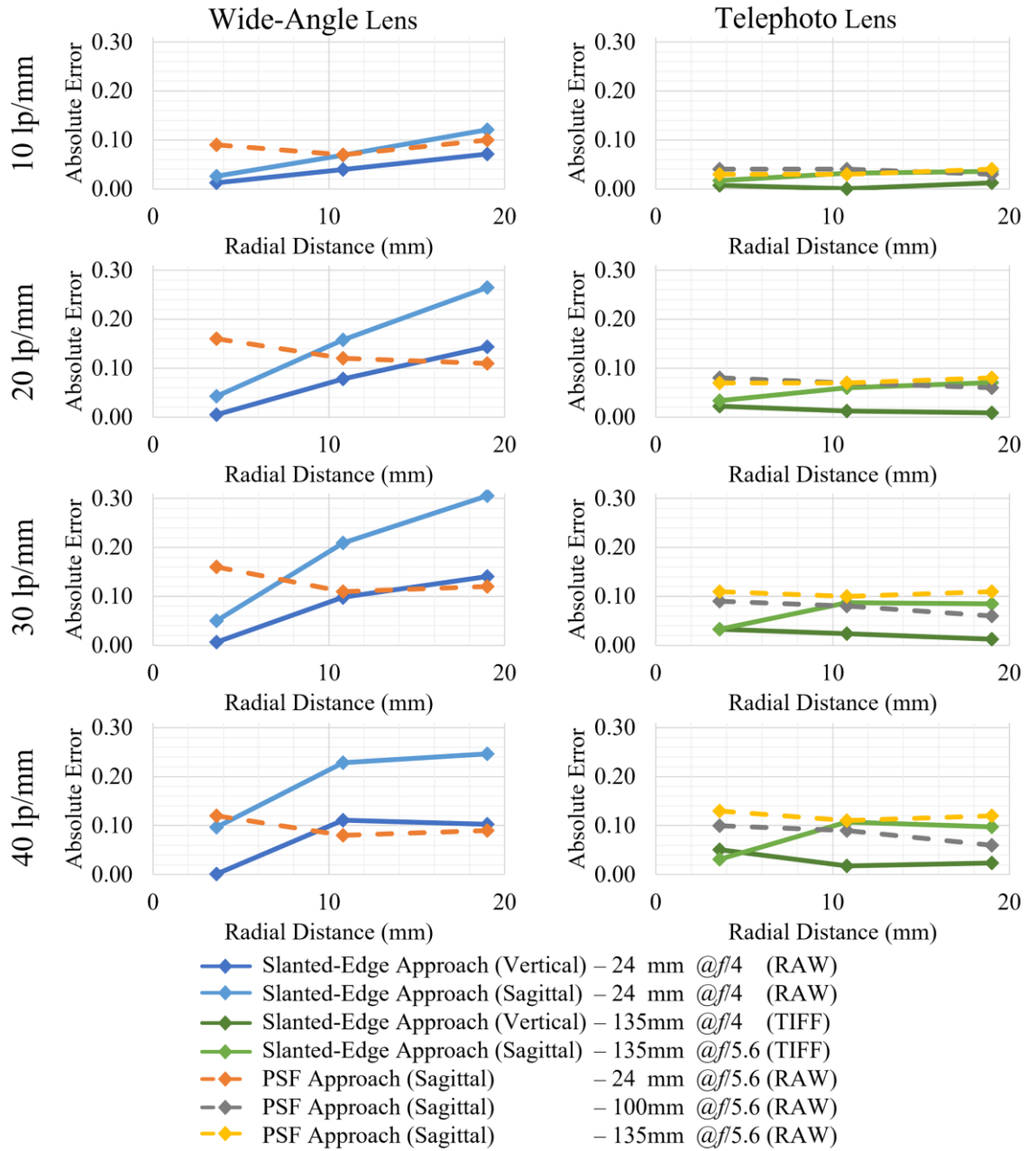


Figure 7.2 The absolute error from the slanted-edge (as presented in this thesis) and PSF-CNN (as published by Bauer et al. [15]) approaches, at 10, 20, 30 and 40 lp/mm for radial distances 1/6, 3/6 and 6/6. This comparison is between similar, but not identical, systems.

7.4 Potential Applications

There is potential utilisation of the proposed method in several areas of imaging science. The following sub-sections address some of these applications, highlighting areas where the method should be adapted for each use case.

7.4.1 Camera Performance Comparison

The comparison of camera system performance is a simple application of the proposed method. A system e-SFR estimate from natural scenes potentially allows multiple systems to be assessed by collecting image datasets of similar content from each system. The camera itself does not necessarily need to be at hand; the images can be harvested from online sources and social media. Large quantities of images can be acquired in such cases, resulting in more representative ISO12233 e-SFR estimates.

Only minor modifications to the method would be required for such comparisons. An image file EXIF reader is needed to extract the camera settings from the metadata, categorising the dataset by model, aperture and focal length, which are essential system performance factors. In addition, the metadata can classify the dataset by ISO gain setting, shutter speed and file format, depending on the variables of interest. Using a similar scene classification CNN as Chapter 6, the datasets could also be sub-classed to determine the performance under various illuminations, scenes, and settings.

It is important to note that many of the images in online sources are often resized, compressed, and stripped of metadata; therefore, careful consideration should be given to deselecting these during image collection.

7.4.2 Image Quality Metrics

As mentioned, the system e-SFR estimation indicated scene dependent characteristics with highly non-linear systems. After implementing the previously proposed study to analyse the scope of this scene dependency (cf. § 7.2.4), the NS-SFRs may be developed into an SPD e-SFR measurement, that in turn, can be used in a similar manner to the SPD-MTF as inputs in spatial IQMs to account for the scene dependent nature on non-linear pipelines [13, 17, 96] (cf. § 2.8.2).

The SPD-MTF and the NS-SFRs are conceptually very different. The SPD e-SFR is based on the slanted-edge method. It will be most sensitive to non-linear sharpening and provide local performance measurement. Whilst the SPD-MTF is a global measurement based on the texture method (cf. § 2.1.4), hence effectively quantify non-linear sharpening and denoising.

When implemented into IQMs, these performance measurements may behave differently. Therefore, further work is required to compare the IQM accuracy of these approaches.

The texture-based SPD-MTF is implemented through printed natural scene test charts [12] or system simulation [13]. Consequently, it cannot obtain the performance from uncharacterised real-world natural scene captures. The NS-SFR are obtained without any former knowledge of the input signal; thus, if SPD e-SFR can be established, they can be implemented into a *no-reference* spatial IQM, opening new possibilities for this approach.

7.4.3 Autonomous Vision Systems

Deep Neural Networks (DNNs) are currently used as one of the leading technologies in image recognition tasks. With the application of DNNs in decision critical systems, such as autonomous vision systems, it is important to develop camera performance measures to monitor the camera output in real-time.

One of the most significant appeals for a performance measure derived from natural scenes is its potential for providing real-time evaluations. A live-SFR feed would benefit several areas of autonomous vision systems, such as autonomous vehicles and security systems.

There are several reasons why camera signal may deteriorate during real-time operation, including camera system failure, motion blur, defocus, and environmental conditions. Through live monitoring, the system can be adapted according to the measured and expected camera performance, adjusting the image signal processing (ISP), or completely removing the automation when the SFR drops under what is deemed safe operation.

The proposed methodology requires significant optimisation to achieve live-SFRs, i.e. an e-SFR estimated per frame, at 30 frames per second ($1/30^{\text{th}}$ of a second). The optimisation would include previously stated code optimisation (cf. § 7.2.1) and application-specific computational improvements. For example, edge detection and isolation using a CNN (cf. § 7.2.1), the application would benefit the CNN being trained with scene contents expected for the purpose. In addition, the methodology can be adapted to isolate edges known to be

expected within specific scene types, such as, in autonomous vehicles, road signs and road markings would be most often in predictable places in the field of view and moving in the frame with the speed and direction of the vehicle, potentially facilitating edge isolation.

Using multiple images (or frames) was shown to increase accuracy and precision. Therefore, an alternative solution would be to estimate the e-SFR every x seconds, combining multiple frames in the estimation. Such an approach would not necessarily be considered live, but would allow autonomous vision systems to be monitored more accurately during operation.

Chapter 8 Conclusions and Recommendations

The conclusion of the research conducted to establish a method to obtain natural scene derived performance measurement and the recommendations for further work are presented in this chapter.

8.1 Conclusions

The following conclusions were drawn from the research conducted in this thesis:

- The thesis presents a novel automated approach to extract appropriate step-edges from natural scenes for processing in the slanted-edge algorithm that resulted in the *natural scene derived spatial frequency response* (NS-SFR).
- A top-down approach was established as the most appropriate method for selecting the natural scene step-edges. First, detecting all image edges using the Canny edge detector and then filtering them down to produce step-edge regions of interest (ROIs).
- The pixel stretching filter was developed to improve the edge extraction from natural scenes. This filter not only isolated the edge of interest successfully, but also reduced the effects of image noise and non-uniformity on the resulting NS-SFR.
- By minimising the variation in the NS-SFR data to improve conformity with the ISO12233 standard method, without excessively restricting the data, the system e-SFR can be estimated within one standard deviation of the equivalent ISO12233 e-SFR for linear camera systems.
- Results from a smartphone system incorporating highly non-linear image signal processing (ISP) indicate that the e-SFR estimate is scene and processing-dependent (SPD). The resultant scene dependency is apparent after classifying the smartphone dataset into man-made, indoor and nature scenes.
- This proposed approach has low NS-SFR data yield, on average, only selecting 3.41% of the collected NS-SFR data for the system e-SFR estimation. Suboptimal scenes or datasets containing few suitable step-edges may lead to missing data in the centre and corners of the field of view.
- Demosaiced image files are adequate for estimating the system e-SFR, provided the ISP is linear or close to linear. Otherwise, the RAW image can be used in the estimation to ensure linearity.

-
- The 10th percentile of the LSF FWHM was the most appropriate threshold to isolate suitable step-edges per radial segment, as it is more likely to exclude the scene content component of the NS-SFR.
 - The use of extensive image datasets (1000-2000 images) from a near-linear camera system provided e-SFR estimates across the field of view comparable to the equivalent ISO12233 standard measurements.
 - The weighted average e-SFR estimate for a near-linear camera system provided results that stay within one standard deviation of the expected field of view distribution of the e-SFRs from the test chart.
 - The sagittal and tangential e-SFR estimates correlated with the e-SFRs measured with test chart edges. However, they had an MAE greater than the standard deviation of the test chart e-SFRs. This error is caused by decreasing the number of isolated step-edges by 71% in the horizontal and vertical to sagittal and tangential orientation conversion. Hence, expanding the LSF FWHM distribution threshold to utilise broader LSFs was shown to reduce this error, as more edges are utilised in the estimation.
 - Extensive image datasets (2000 images) from a highly non-linear camera system produced e-SFR estimates that indicated scene and processing dependant behaviour. This was further demonstrated when subdividing the dataset into three scene classes (man-made, indoor and nature) that induced different levels of processing.
 - Subdividing the DSLR 1 dataset into smaller subsets provided important observations on the accuracy and precision of the system e-SFR estimate:
 - The larger image subsets increase the accuracy and precision for radial distance segments with low numbers of isolated step-edges, such as in the centre segment.
 - Larger subsets can decrease estimation accuracy for radial segments containing optical distortion artifacts and large e-SFR variation, such as the frame corners. This is due to artifacts increasing the spread of the LSF and taking an average of a large optical performance variation. Whilst smaller subsets will sample narrow parts of the e-SFR variation, increasing accuracy, but decreasing precision.
 - Partway radial distances were shown to contain 80% more isolated step-edges than the other radial segments; these areas maintained the same average accuracy across the subset sizes and all scene classes, but the precision increased with larger datasets.

-
- The weighted average gave the most stable result for small subsets, staying within one ISO12233 e-SFR standard deviation, but accuracy and precision improve using larger subsets.
 - Captured scenes of man-made and indoor structures contain many isolated step-edges, but indoor scenes contain more image noise due to the lower illumination condition. Scenes of nature are generally more busy, with natural scene textures and few easily isolatable step-edges. As expected, the scenes containing more suitable step-edges result in more accurate and precise system e-SFRs estimates with fewer images.
 - The proposed method has the potential to improve several areas of imaging science, including:
 - Camera performance comparison using online image datasets.
 - Produce a SPD SFR that may open up new possibilities in *no-reference* spatial IQM.
 - Create a real-time SFR measurement for monitoring/calibrating autonomous vision systems.

8.2 Further Work

The further work recommendations are split into two parts; the first presents improvements in the methodology proposed in this thesis, and the second lists other related work.

Further work recommendations for the proposed methodology:

- The first recommended development is to optimise the proposed method to improve the user application and computation times, enabling it to be practically implemented. This optimisation should improve the NS-SFR yield in the system e-SFR estimation by initially selecting the desired step-edge parameters, removing approximately 66% of unprocessed data. Additional optimisation includes hardware acceleration, reducing the computation time.
- Training a Convolutional Neural Network (CNN) to optimise the edge selection and verification would significantly improve computation times. Rather than taking the Canny edge detection top-down approach (cf. § 3.2.1), this would be more comparable to the matched filter (cf. § 3.2.2). Instead of building up the appropriate edges, the CNN

could be trained using the data gathered by the current approach, selecting the suitable edges straight from the scenes. This method would still benefit from isolation methods such as pixel stretching (cf. § 3.3.2).

- The pixel stretching filter could be replaced using the Edge Spread Function (ESF) mask as a Hamming filter in the ISO12233 algorithm to improve computation time further. Research should be carried out into the effects of this approach on the e-SFR and whether it would remain comparable to the standard.
- Parameters such as the optical focus distance should be implemented into the system e-SFR estimation. A possible solution would be to estimate a single image depth map and combine it with the location of the highest performance NS-SFRs (i.e. the narrowest LSFs). That said, further research is required into the field of monocular depth estimation before such an approach can be successfully implemented (cf. § 2.9).
- Motion blur is another parameter that has not been explored in the course of this project. Thresholding the LSF FWHM removed blurred edges; thus, motion blur is not a concern. However, further work to account for motion blur in situations where the Inertial Measurement Unit (IMU) is known (cf. § 2.9) may benefit specific applications.
- Although pixel stretching has been shown to mitigate the effects of tonal non-uniformity, it does not entirely remove it. Therefore, further work to develop an alternative solution to remove non-uniformity would enhance the proposed method, for instance, describing the tonal changes in the form of a function and applying the inverse.
- Applying the latest improvements to the ISO12233 standard, particularly research proposals that have shown to help reduce variance in the e-SFR, such as the Reverse-Projection method [64], would benefit the proposed method.

Further recommendations for the other related works:

- The SPD of the NS-SFRs and the resulting system e-SFR estimates should be further researched to establish how the results are affected by non-linear ISPs on various scene types, both using global and local image content. Such a study would provide insight into whether such an approach would be a more suitable non-linear system measurement than the traditional laboratory-based ISO12233 e-SFR.
- If the resulting SPD SFR measurement is effective, research should be conducted into whether it can substitute the SPD-Modulation Transfer Functions (SPD-MTF) in spatial

image quality metrics (IQM) [13, 17]. This work should determine the error and IQM accuracy of the two very conceptually different methods.

- The performance of the proposed slanted-edge approach should be compared to alternative solutions, such as the PSF-CNN approach [15] (cf. § 2.8.4 & 7.3). This comparison should use independent datasets, benchmarking the accuracy with various ISPs and dataset sizes.
- Finally, the proposed method should be adapted and implemented for specific practical applications, evaluating its effectiveness for various purposes. For example, estimating the system e-SFR from datasets for Deep Neural Networks DNN training, security systems, and autonomous driving.

Appendix A The ISO12233 e-SFR Measurement

Figure A.1 contains the Opto-Electric Conversion Function (OECF) of the *Nikon D800* DSLR camera system, measured using the ISO14524:2009 [2]. This OECF was used in the linearisation process for the *Nikon D800* ISO12233 system e-SFR measurements.

Figure A.2 contains the average system e-SFR for the DSLR1 system (cf. § 5.3.1), measured from a TIFF pipeline (demosaiced) and RAW pipeline (green channel mosaiced) across six radial distance segments and a weighted average of the entire frame. They were measured following the ISO12233:2017 standard [5]. There is a negligible difference between the TIFF and RAW pipelines; hence, all DSLR 1 system e-SFR estimations were evaluated against the ISO12233 system e-SFR measured from TIFF images.

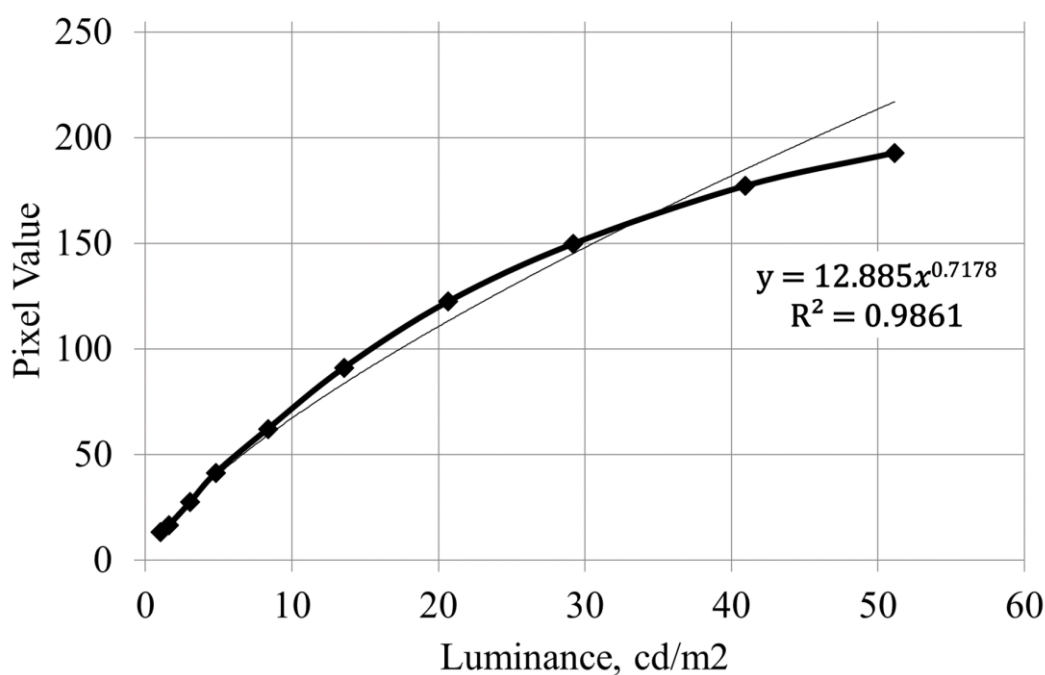


Figure A.1 The OECF of the *Nikon D800*, measured using the ISO14524:2009 [2] standard method.

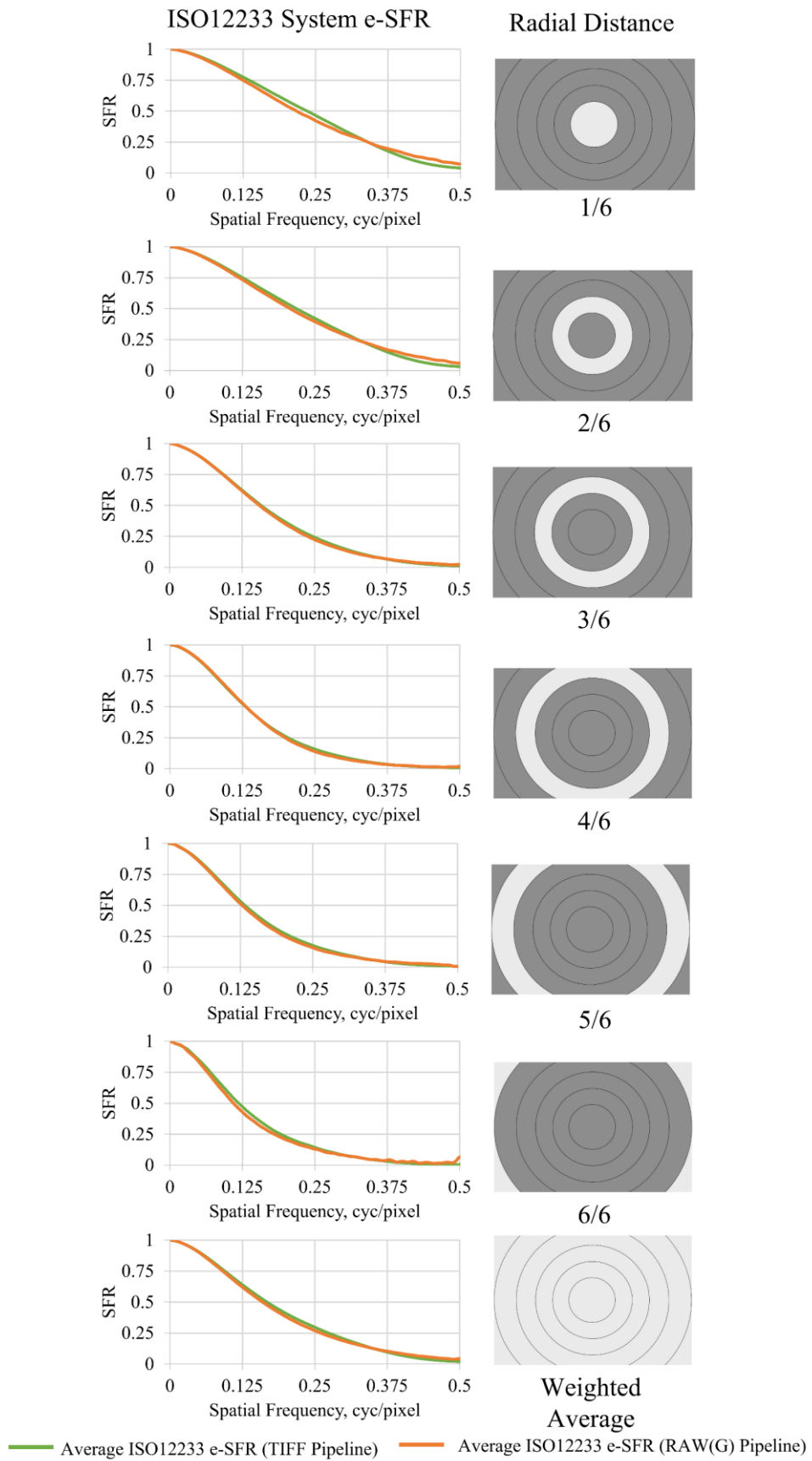


Figure A.2 Comparison between the system e-SFR from the DSLR 1 TIFF and RAW pipelines, measured using the ISO12233:2017 [5] standard method.

Appendix B E-SFR Estimation Summary Metric

Tables B.1 and B.2 show the e-SFR summary metric of the vertical system estimates from DSLR 1 TIFF and RAW datasets, respectively. Table B.3 contains the same summary metric for the DSLR 2 dataset and Table B.4 for the smartphone system. Metrics used were the MTF50/50P/20/10, 25%/50%/75% of the Nyquist frequency and sampling efficiency as detailed in Chapter 2 (cf. § 2.6.1), as well as LSF FWHM (cf. § 5.1). These metrics were calculated for each radial segment system e-SFR estimation and weighted mean. The absolute error was calculated for each metric, respective to equivalent ISO12233 e-SFR. These error values were colour coded to visualise whether the summary metrics are equal to ISO12233 (green), overestimated (blue) or underestimated (red).

Measurement	Radial Distance												Weighted Ave	
	1		2		3		4		5		6		Value	Err
	Value	Err	Value	Err	Value	Err	Value	Err	Value	Err	Value	Err	Value	Err
10th Percentile of LSF FWHM Distribution														
MTF50	0.24	0.01	0.23	0.01	0.21	0.05	0.19	0.06	0.20	0.06	0.16	0.04	0.21	0.04
MTF50P	0.24	0.01	0.23	0.01	0.21	0.05	0.19	0.06	0.20	0.06	0.16	0.04	0.21	0.04
MTF20	0.38	0.01	0.37	0.02	0.34	0.07	0.33	0.10	0.33	0.09	0.29	0.07	0.35	0.05
MTF10	0.44	0.02	0.45	0.04	0.41	0.07	0.40	0.11	0.41	0.10	0.38	0.10	0.42	0.05
25% Nyquist freq.	0.82	0.04	0.78	0.03	0.77	0.14	0.72	0.20	0.73	0.20	0.63	0.16	0.75	0.11
50% Nyquist freq.	0.48	0.02	0.44	0.02	0.39	0.15	0.35	0.19	0.36	0.18	0.26	0.12	0.39	0.10
75% Nyquist freq.	0.20	0.02	0.20	0.05	0.15	0.08	0.13	0.09	0.14	0.09	0.11	0.08	0.16	0.06
Sampling Efficiency	88.60	4.51	90.22	8.84	82.73	14.76	80.33	21.56	81.56	20.07	76.90	20.29	84.78	10.53
LSF FWHM*	5.00	-1.00	4.00	-2.00	7.00	-1.00	7.00	-3.00	7.00	-3.00	8.00	-3.00	7.00	-1.00
20th Percentile of LSF FWHM Distribution														
MTF50	0.24	0.00	0.22	0.00	0.19	0.03	0.17	0.04	0.18	0.05	0.16	0.04	0.20	0.03
MTF50P	0.24	0.00	0.22	0.00	0.19	0.03	0.17	0.04	0.18	0.05	0.16	0.04	0.20	0.03
MTF20	0.37	0.01	0.36	0.01	0.31	0.04	0.29	0.07	0.30	0.07	0.27	0.05	0.33	0.03
MTF10	0.44	0.02	0.43	0.03	0.38	0.04	0.37	0.07	0.38	0.07	0.36	0.08	0.41	0.04
25% Nyquist freq.	0.81	0.04	0.78	0.03	0.72	0.09	0.67	0.15	0.70	0.17	0.62	0.14	0.73	0.09
50% Nyquist freq.	0.47	0.01	0.42	0.00	0.32	0.08	0.28	0.12	0.31	0.13	0.24	0.09	0.36	0.06
75% Nyquist freq.	0.19	0.02	0.17	0.03	0.11	0.04	0.09	0.06	0.11	0.06	0.09	0.06	0.13	0.04
Sampling Efficiency	87.99	3.90	86.97	5.59	76.49	8.53	73.55	14.78	76.30	14.81	71.72	15.10	81.34	7.08
LSF FWHM*	6.00	0.00	6.00	0.00	8.00	0.00	8.00	-2.00	8.00	-2.00	8.00	-3.00	7.00	-1.00
50th Percentile of LSF FWHM Distribution														
MTF50	0.21	-0.02	0.18	-0.03	0.16	0.00	0.14	0.01	0.15	0.02	0.13	0.01	0.17	0.00
MTF50P	0.21	-0.02	0.18	-0.03	0.16	0.00	0.14	0.01	0.15	0.02	0.13	0.01	0.17	0.00
MTF20	0.34	-0.02	0.31	-0.04	0.26	-0.01	0.24	0.02	0.25	0.02	0.22	0.01	0.28	-0.02
MTF10	0.41	-0.01	0.38	-0.03	0.33	-0.01	0.31	0.02	0.32	0.01	0.29	0.01	0.36	-0.01
25% Nyquist freq.	0.77	-0.01	0.71	-0.05	0.64	0.01	0.58	0.06	0.61	0.08	0.53	0.06	0.66	0.02
50% Nyquist freq.	0.39	-0.07	0.32	-0.11	0.23	-0.01	0.19	0.03	0.21	0.03	0.15	0.01	0.27	-0.03
75% Nyquist freq.	0.14	-0.03	0.11	-0.04	0.06	0.00	0.05	0.02	0.06	0.01	0.05	0.02	0.09	-0.01
Sampling Efficiency	84.09	0.00	81.38	0.00	67.96	0.00	58.77	0.00	61.49	0.00	56.61	0.00	74.25	0.00
LSF FWHM*	8.00	2.00	8.00	2.00	8.00	0.00	8.00	-2.00	8.00	-2.00	8.00	-3.00	8.00	0.00

* LSF FWHM is measured from the 4-times resampled edge, thus, 4-times the native resolution

Table B.1 Metrics for the vertical system e-SFR estimates from DSLR 1 TIFF dataset. Eight metrics for each radial segment and the weighted average per LSF FWHM percentile threshold have been measured.

Measurement	Radial Distance												Weighted Ave	
	1		2		3		4		5		6		Value	Err
10th Percentile of LSF FWHM Distribution														
MTF50	0.22	0.00	0.20	0.00	0.19	0.03	0.15	0.02	0.17	0.04	0.14	0.03	0.18	0.02
MTF50P	0.22	0.00	0.20	0.00	0.19	0.03	0.15	0.02	0.17	0.04	0.14	0.03	0.18	0.02
MTF20	0.41	0.04	0.47	0.12	0.32	0.06	0.43	0.21	0.31	0.09	0.25	0.05	0.32	0.03
MTF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25% Nyquist freq.	0.75	0.00	0.74	0.01	0.71	0.09	0.60	0.07	0.64	0.13	0.58	0.15	0.68	0.06
50% Nyquist freq.	0.42	0.00	0.39	0.00	0.33	0.11	0.20	0.06	0.31	0.15	0.20	0.07	0.32	0.05
75% Nyquist freq.	0.22	0.02	0.21	0.04	0.14	0.07	0.10	0.07	0.14	0.09	0.09	0.05	0.15	0.05
Sampling Efficiency	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LSF FWHM*	2.00	0.00	2.00	-1.00	2.00	-2.00	3.00	-2.00	2.00	-2.00	4.00	-1.00	3.00	-1.00
20th Percentile of LSF FWHM Distribution														
MTF50	0.20	-0.01	0.17	-0.03	0.15	0.00	0.14	0.01	0.15	0.02	0.13	0.02	0.16	0.00
MTF50P	0.20	-0.01	0.17	-0.03	0.15	0.00	0.14	0.01	0.15	0.02	0.13	0.02	0.16	0.00
MTF20	0.35	-0.01	0.29	-0.05	0.27	0.02	0.23	0.03	0.30	0.08	0.24	0.02	0.28	0.00
MTF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25% Nyquist freq.	0.72	-0.02	0.67	-0.05	0.61	0.02	0.56	0.06	0.59	0.09	0.53	0.07	0.62	0.02
50% Nyquist freq.	0.40	-0.01	0.27	-0.10	0.24	0.04	0.17	0.06	0.23	0.09	0.18	0.02	0.26	0.01
75% Nyquist freq.	0.19	-0.01	0.18	0.01	0.11	0.04	0.10	0.07	0.12	0.07	0.08	0.04	0.13	0.03
Sampling Efficiency	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LSF FWHM*	2.00	0.00	2.00	-1.00	2.00	-2.00	3.00	-2.00	2.00	-2.00	4.00	-1.00	3.00	-1.00
50th Percentile of LSF FWHM Distribution														
MTF50	0.17	-0.04	0.15	-0.05	0.12	-0.02	0.12	-0.01	0.12	0.00	0.12	0.00	0.14	-0.02
MTF50P	0.17	-0.04	0.15	-0.05	0.12	-0.02	0.12	-0.01	0.12	0.00	0.12	0.00	0.14	-0.02
MTF20	0.30	-0.06	0.26	-0.08	0.21	-0.04	0.25	0.04	0.22	0.01	0.20	-0.02	0.24	-0.04
MTF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25% Nyquist freq.	0.66	-0.08	0.61	-0.10	0.50	-0.09	0.46	-0.03	0.49	-0.01	0.46	0.00	0.54	-0.06
50% Nyquist freq.	0.29	-0.11	0.22	-0.15	0.15	-0.05	0.10	-0.01	0.15	0.01	0.12	-0.05	0.18	-0.07
75% Nyquist freq.	0.14	-0.06	0.11	-0.05	0.08	0.01	0.07	0.03	0.08	0.04	0.05	0.01	0.09	-0.01
Sampling Efficiency	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LSF FWHM*	2.00	0.00	2.00	-1.00	4.00	0.00	4.00	-1.00	4.00	0.00	4.00	-1.00	4.00	0.00

* LSF *FWHM* is measured from the 4-times resampled edge, thus, 4-times the native resolution

Table B.2 Metrics for the vertical system e-SFR estimates from DSLR 1 RAW dataset. Eight metrics for each radial segment and the weighted average per LSF FWHM percentile threshold have been measured.

Measurement	Radial Distance												Weighted Ave	
	1		2		3		4		5		6		Value	Err
10th Percentile of LSF FWHM Distribution														
MTF50	0.21	-0.02	0.21	-0.02	0.21	0.00	0.20	-0.01	0.19	-0.01	0.19	-0.01	0.20	-0.01
MTF50P	0.21	-0.02	0.21	-0.02	0.21	0.00	0.20	-0.01	0.19	-0.01	0.19	-0.01	0.20	-0.01
MTF20	0.35	-0.01	0.35	-0.01	0.35	0.01	0.33	0.01	0.32	0.01	0.32	0.00	0.34	0.00
MTF10	0.42	0.01	0.42	0.01	0.42	0.02	0.40	0.02	0.39	0.02	0.39	0.01	0.41	0.01
25% Nyquist freq.	0.76	-0.03	0.74	-0.03	0.74	-0.02	0.73	0.00	0.72	-0.01	0.73	0.00	0.74	-0.02
50% Nyquist freq.	0.41	-0.06	0.40	-0.05	0.41	-0.01	0.36	-0.01	0.35	0.00	0.35	-0.03	0.38	-0.04
75% Nyquist freq.	0.15	-0.01	0.16	0.00	0.16	0.03	0.13	0.02	0.12	0.03	0.11	0.01	0.14	0.01
Sampling Efficiency	84.21	1.12	84.14	2.28	83.72	4.20	79.85	4.12	78.96	4.57	77.44	1.15	81.47	2.09
LSF FWHM*	2.00	-4.00	6.00	0.00	7.00	0.00	7.00	-1.00	8.00	0.00	8.00	0.00	7.00	0.00
20th Percentile of LSF FWHM Distribution														
MTF50	0.20	-0.04	0.19	-0.04	0.19	-0.03	0.17	-0.03	0.17	-0.03	0.17	-0.03	0.18	-0.03
MTF50P	0.20	-0.04	0.19	-0.04	0.19	-0.03	0.17	-0.03	0.17	-0.03	0.17	-0.03	0.18	-0.03
MTF20	0.32	-0.04	0.32	-0.04	0.31	-0.03	0.29	-0.03	0.29	-0.02	0.29	-0.04	0.31	-0.03
MTF10	0.39	-0.02	0.39	-0.02	0.38	-0.02	0.37	-0.01	0.36	-0.01	0.36	-0.02	0.38	-0.02
25% Nyquist freq.	0.74	-0.05	0.70	-0.07	0.70	-0.06	0.68	-0.06	0.67	-0.06	0.68	-0.05	0.70	-0.06
50% Nyquist freq.	0.36	-0.11	0.34	-0.12	0.33	-0.09	0.29	-0.08	0.28	-0.07	0.27	-0.10	0.32	-0.10
75% Nyquist freq.	0.12	-0.05	0.12	-0.04	0.11	-0.03	0.09	-0.01	0.09	-0.01	0.09	-0.02	0.10	-0.03
Sampling Efficiency	78.22	-4.87	77.91	-3.95	76.14	-3.38	73.30	-2.43	72.32	-2.07	72.35	-3.93	75.49	-3.89
LSF FWHM*	8.00	2.00	8.00	2.00	8.00	1.00	8.00	0.00	8.00	0.00	8.00	0.00	8.00	1.00
50th Percentile of LSF FWHM Distribution														
MTF50	0.15	-0.09	0.14	-0.10	0.14	-0.08	0.14	-0.06	0.13	-0.06	0.14	-0.07	0.14	-0.08
MTF50P	0.15	-0.09	0.14	-0.10	0.14	-0.08	0.14	-0.06	0.13	-0.06	0.14	-0.07	0.14	-0.08
MTF20	0.26	-0.10	0.24	-0.11	0.24	-0.10	0.24	-0.08	0.23	-0.08	0.23	-0.09	0.24	-0.10
MTF10	0.33	-0.08	0.32	-0.09	0.31	-0.09	0.31	-0.06	0.30	-0.07	0.31	-0.07	0.31	-0.08
25% Nyquist freq.	0.60	-0.18	0.55	-0.23	0.55	-0.21	0.57	-0.17	0.54	-0.18	0.55	-0.19	0.56	-0.20
50% Nyquist freq.	0.22	-0.25	0.19	-0.27	0.18	-0.24	0.19	-0.18	0.17	-0.18	0.17	-0.20	0.19	-0.23
75% Nyquist freq.	0.06	-0.11	0.06	-0.10	0.05	-0.08	0.06	-0.05	0.04	-0.05	0.05	-0.06	0.05	-0.08
Sampling Efficiency	66.38	-16.71	64.56	-17.30	62.05	-17.48	62.91	-12.81	59.87	-14.51	61.68	-14.60	62.99	-16.38
LSF FWHM*	8.00	2.00	8.00	2.00	8.00	1.00	8.00	0.00	8.00	0.00	8.00	0.00	8.00	1.00

* LSF HPW is measured from the 4-times resampled edge, thus, 4-times the native resolution

Table B.3 Metrics for the vertical system e-SFR estimates from DSLR 2 TIFF dataset. Eight metrics for each radial segment and the weighted average per LSF FWHM percentile threshold have been measured.

Measurement	Radial Distance												Weighted Ave	
	1		2		3		4		5		6		Value	Err
10th Percentile of LSF FWHM Distribution														
MTF50	0.26	-0.06	0.28	0.02	0.28	0.05	0.23	0.00	0.23	-0.02	0.20	-0.05	0.25	-0.01
MTF50P	0.26	-0.04	0.28	0.03	0.28	0.06	0.23	0.00	0.23	-0.01	0.20	-0.04	0.25	0.00
MTF20	0.39	-0.04	0.42	0.06	0.41	0.08	0.35	0.01	0.36	0.01	0.30	-0.05	0.38	0.02
MTF10	0.51	0.02	0.55	0.14	0.52	0.13	0.45	0.06	0.45	0.03	0.45	0.03	0.49	0.07
25% Nyquist freq.	0.89	-0.22	0.92	-0.04	0.89	0.01	0.84	-0.09	0.84	-0.12	0.76	-0.13	0.86	-0.10
50% Nyquist freq.	0.54	-0.25	0.60	0.06	0.59	0.16	0.45	-0.01	0.45	-0.05	0.32	-0.16	0.51	-0.04
75% Nyquist freq.	0.23	-0.11	0.27	0.11	0.26	0.14	0.17	0.03	0.18	0.01	0.14	-0.02	0.21	0.03
Sampling Efficiency	102.41	4.51	109.99	27.84	103.98	26.47	90.98	11.13	90.80	6.46	89.98	5.97	98.34	13.35
LSF FWHM*	2.00	-4.00	2.00	-6.00	2.00	-6.00	2.00	-5.00	2.00	-5.00	5.00	-3.00	2.00	-5.00
20th Percentile of LSF FWHM Distribution														
MTF50	0.26	-0.07	0.25	-0.01	0.25	0.02	0.22	-0.01	0.21	-0.04	0.18	-0.06	0.23	-0.03
MTF50P	0.26	-0.05	0.25	0.00	0.25	0.02	0.22	-0.01	0.21	-0.03	0.18	-0.05	0.23	-0.02
MTF20	0.38	-0.05	0.38	0.02	0.37	0.04	0.34	0.00	0.33	-0.03	0.30	-0.06	0.36	-0.01
MTF10	0.49	0.00	0.49	0.08	0.47	0.08	0.43	0.03	0.41	-0.01	0.43	0.01	0.45	0.03
25% Nyquist freq.	0.88	-0.23	0.86	-0.10	0.85	-0.03	0.81	-0.12	0.78	-0.17	0.71	-0.18	0.82	-0.14
50% Nyquist freq.	0.52	-0.27	0.51	-0.03	0.50	0.06	0.42	-0.04	0.38	-0.13	0.29	-0.19	0.45	-0.10
75% Nyquist freq.	0.21	-0.12	0.21	0.05	0.20	0.08	0.15	0.02	0.14	-0.03	0.13	-0.04	0.17	-0.01
Sampling Efficiency	97.90	0.00	97.71	15.56	94.29	16.78	85.99	6.14	82.38	-1.95	85.31	1.30	90.15	5.16
LSF FWHM*	2.00	-4.00	2.00	-6.00	2.00	-6.00	2.00	-5.00	2.00	-5.00	5.00	-3.00	2.00	-5.00
50th Percentile of LSF FWHM Distribution														
MTF50	0.22	-0.10	0.21	-0.05	0.21	-0.02	0.19	-0.05	0.18	-0.07	0.16	-0.08	0.20	-0.06
MTF50P	0.22	-0.08	0.21	-0.04	0.21	-0.02	0.19	-0.04	0.18	-0.06	0.16	-0.07	0.20	-0.05
MTF20	0.34	-0.08	0.33	-0.03	0.33	0.00	0.31	-0.03	0.30	-0.06	0.26	-0.09	0.32	-0.05
MTF10	0.42	-0.07	0.41	0.00	0.41	0.02	0.38	-0.02	0.37	-0.05	0.34	-0.08	0.39	-0.03
25% Nyquist freq.	0.81	-0.30	0.77	-0.18	0.77	-0.11	0.73	-0.19	0.72	-0.24	0.66	-0.24	0.75	-0.21
50% Nyquist freq.	0.42	-0.38	0.38	-0.16	0.38	-0.05	0.33	-0.13	0.29	-0.21	0.22	-0.26	0.35	-0.20
75% Nyquist freq.	0.15	-0.18	0.14	-0.03	0.13	0.02	0.10	-0.03	0.10	-0.07	0.08	-0.08	0.12	-0.06
Sampling Efficiency	84.39	-13.51	81.98	-0.17	81.73	4.22	75.92	-3.92	74.48	-9.86	67.37	-16.64	78.97	-6.02
LSF FWHM*	2.00	-4.00	5.00	-3.00	2.00	-6.00	5.00	-2.00	8.00	1.00	8.00	0.00	5.00	-2.00

* LSF HPW is measured from the 4-times resampled edge, thus, 4-times the native resolution

Table B.4 Metrics for the vertical system e-SFR estimates from smartphone JPEG dataset. Eight metrics for each radial segment and the weighted average per LSF FWHM percentile threshold have been measured.

Appendix C All Vertical e-SFR Estimation Results

Figures C.1 to C.9 provide the system e-SFR estimations of the DSLR 1, DSLR 2 and smartphone systems across all six radial segments used. Three system e-SFR estimations are provided per system, using the three LSF FWHM distribution thresholds (10th, 20th and 50th percentiles).

C.1 DSLR 1 e-SFR Estimation Results

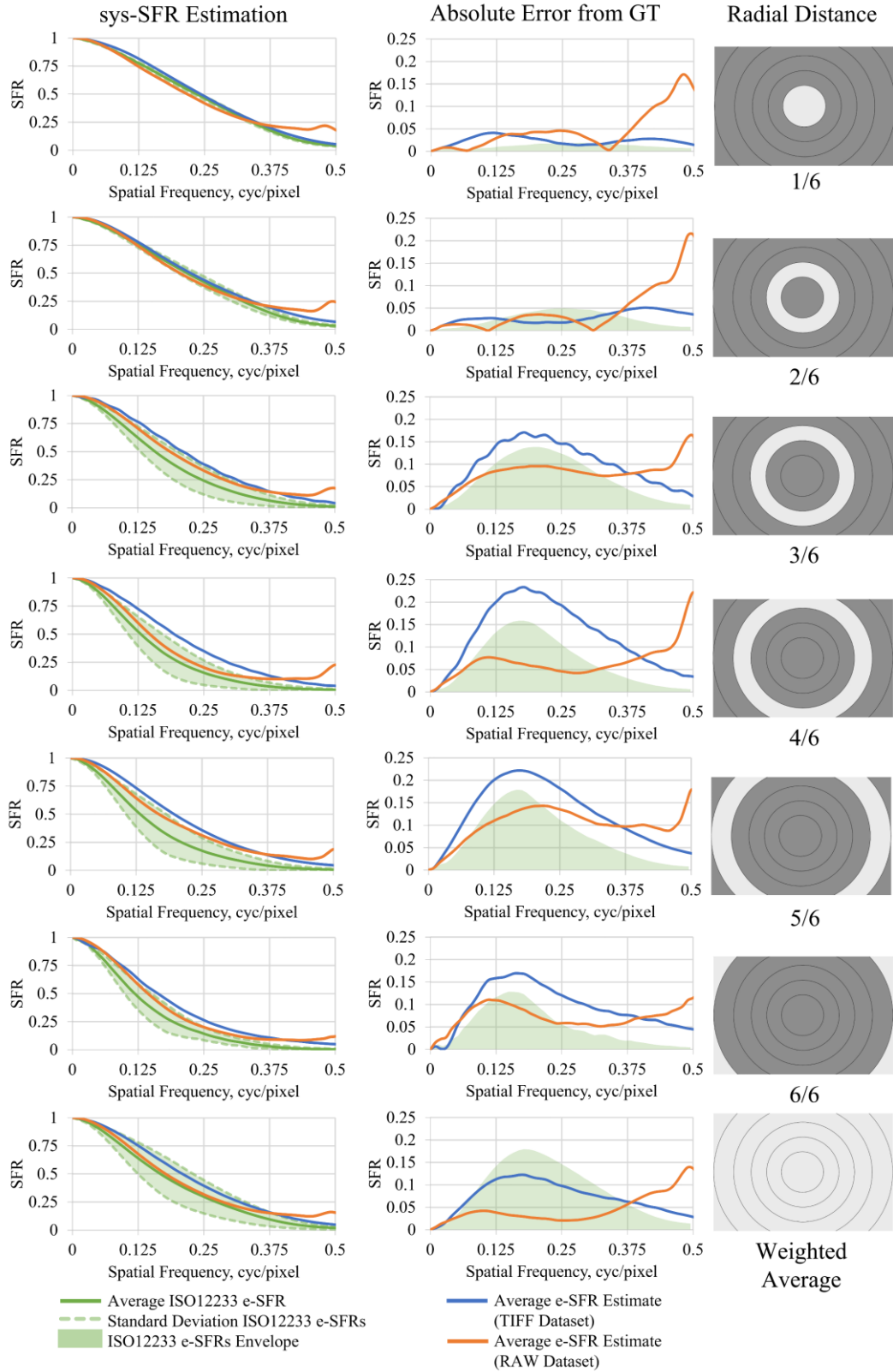


Figure C.1 DSLR 1 vertical system e-SFR estimations using the 10th percentile of the LSF FWHM distribution.

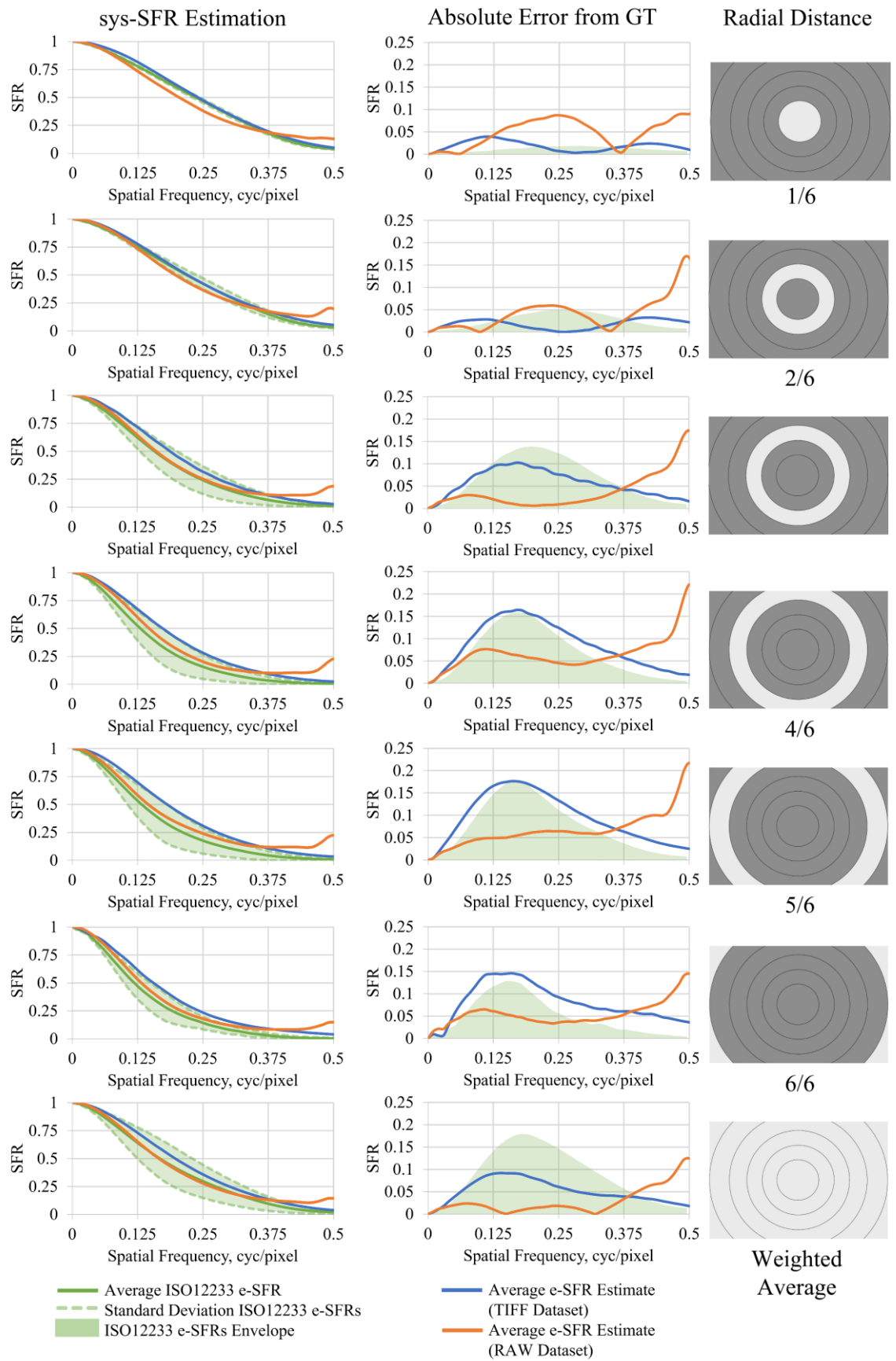


Figure C.2 DSLR 1 vertical system e-SFR estimations using the 20th percentile of the LSF FWHM distribution.

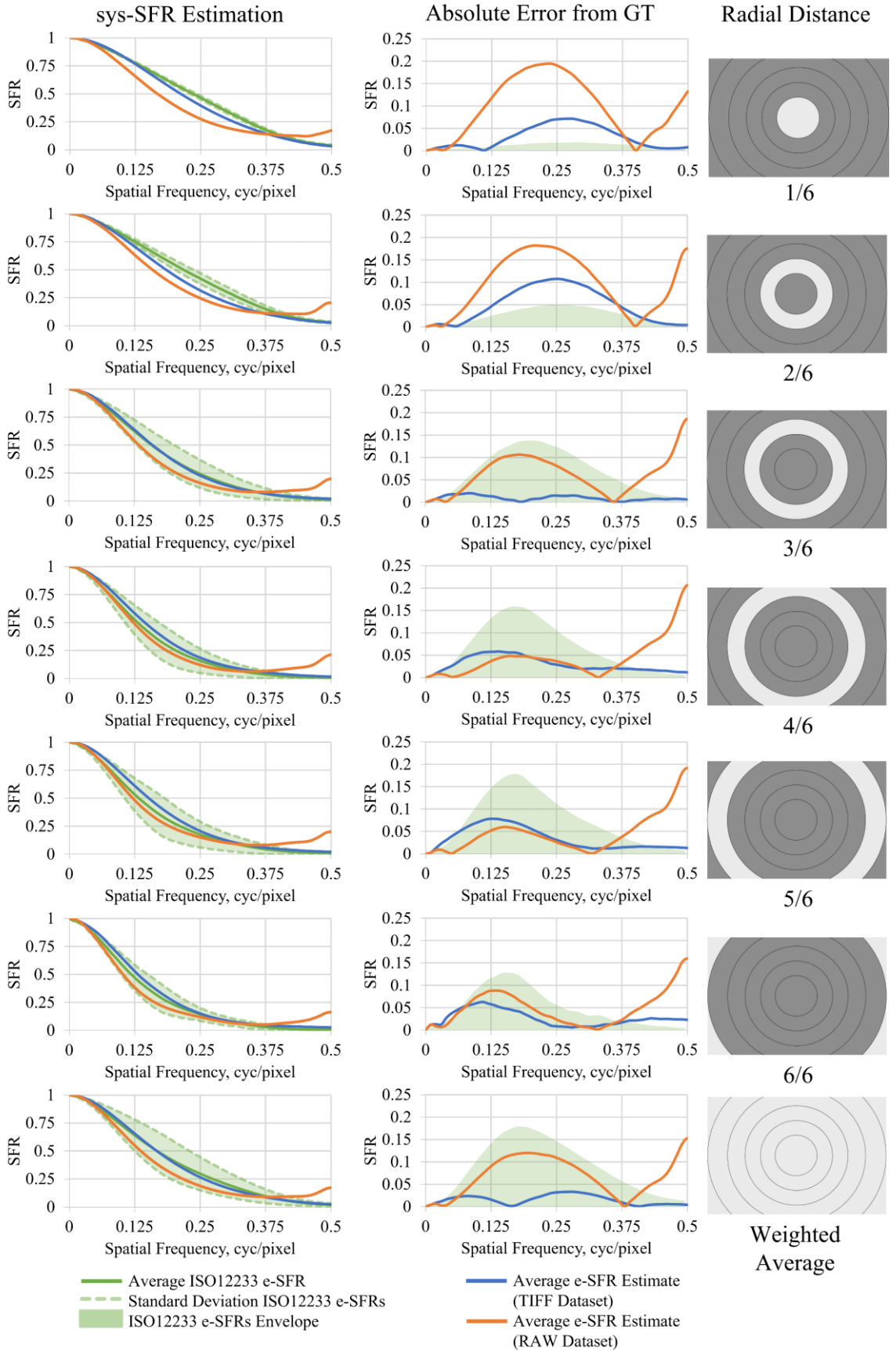


Figure C.3 DSLR 1 vertical system e-SFR estimations using the 50th percentile of the LSF FWHM distribution.

C.2 DSLR 2 e-SFR Estimation Results

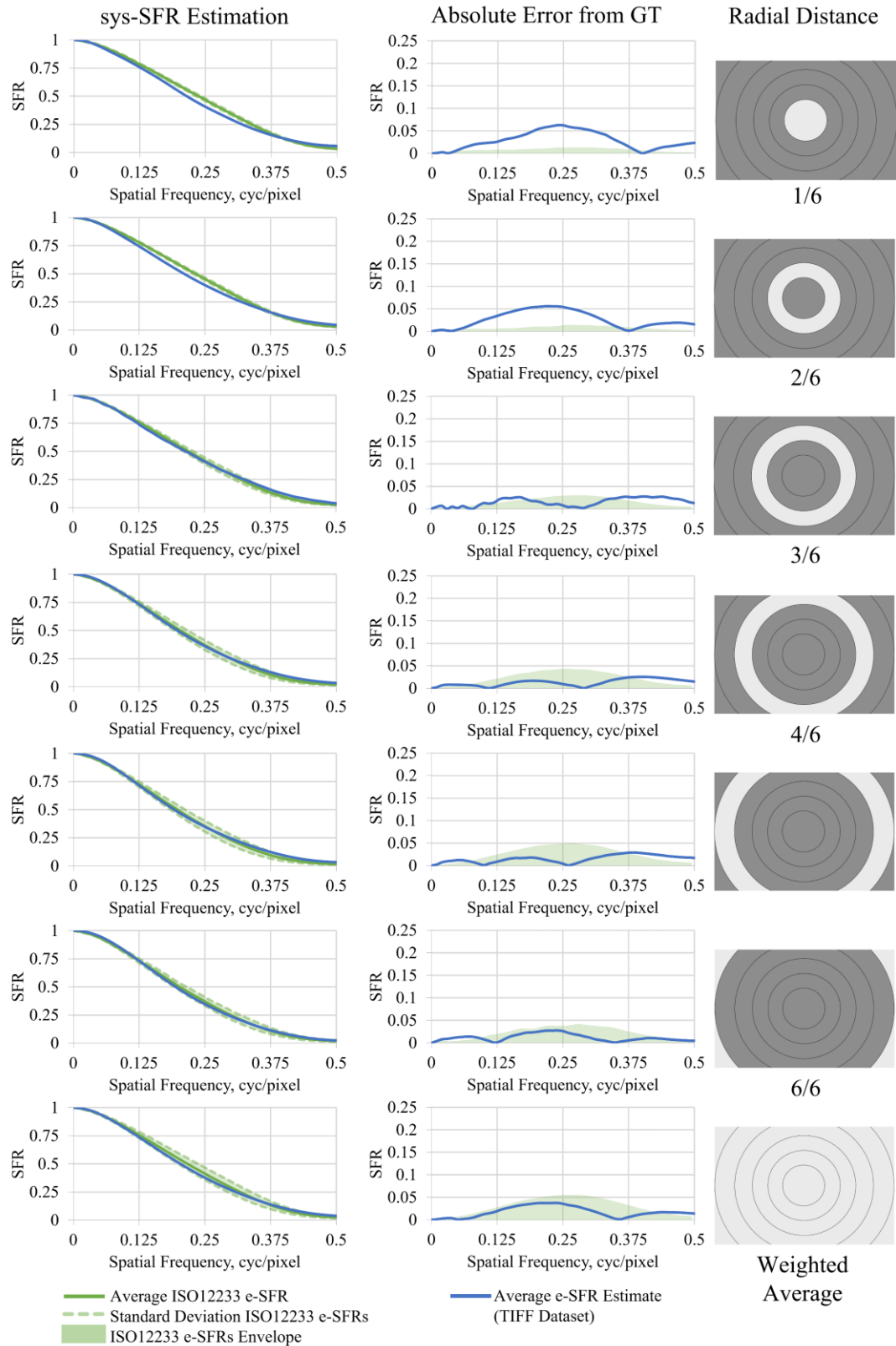


Figure C.4 DSLR 2 vertical system e-SFR estimations using the 10th percentile of the LSF FWHM distribution

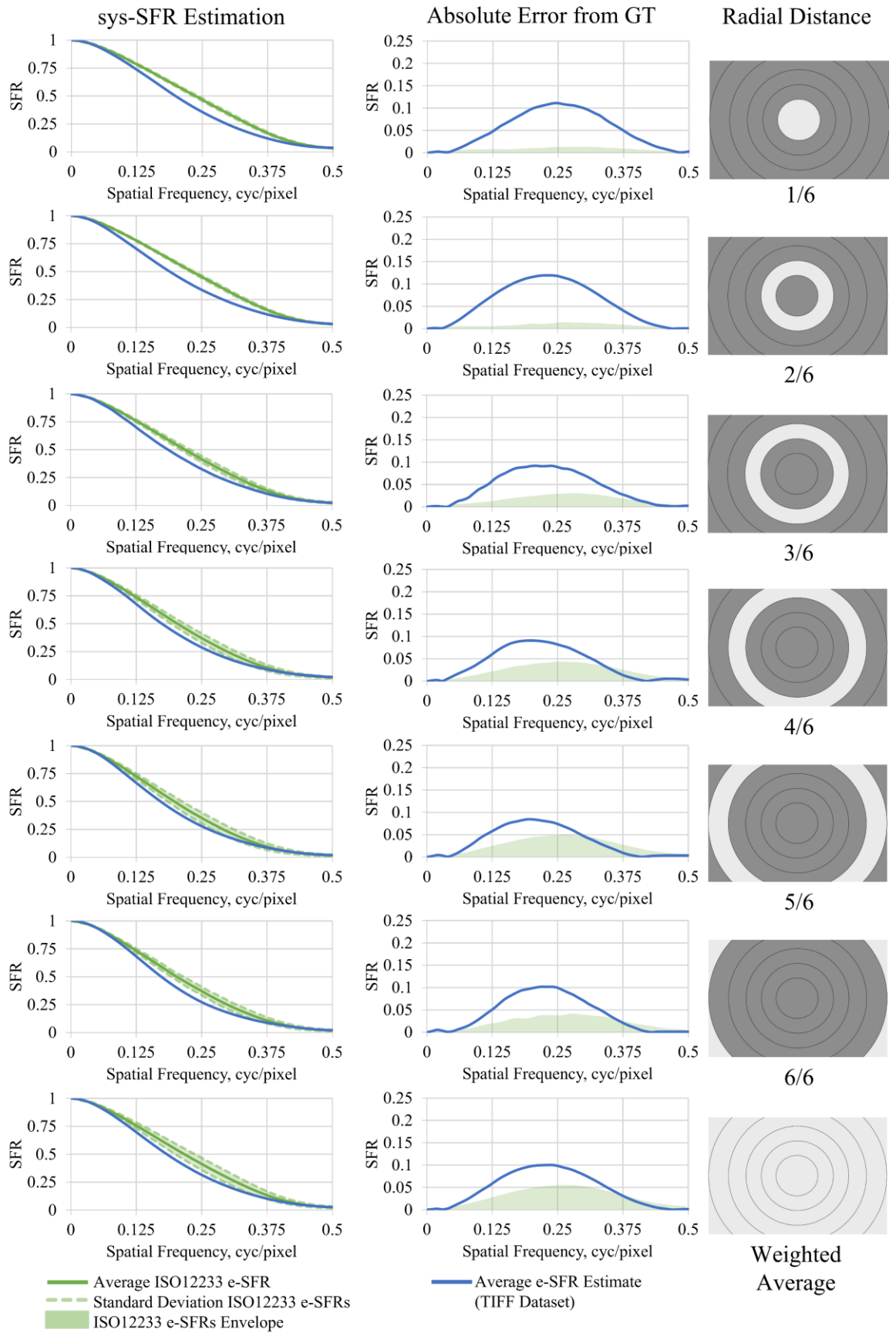


Figure C.5 DSLR 2 vertical system e-SFR estimations using the 20th percentile of the LSF FWHM distribution

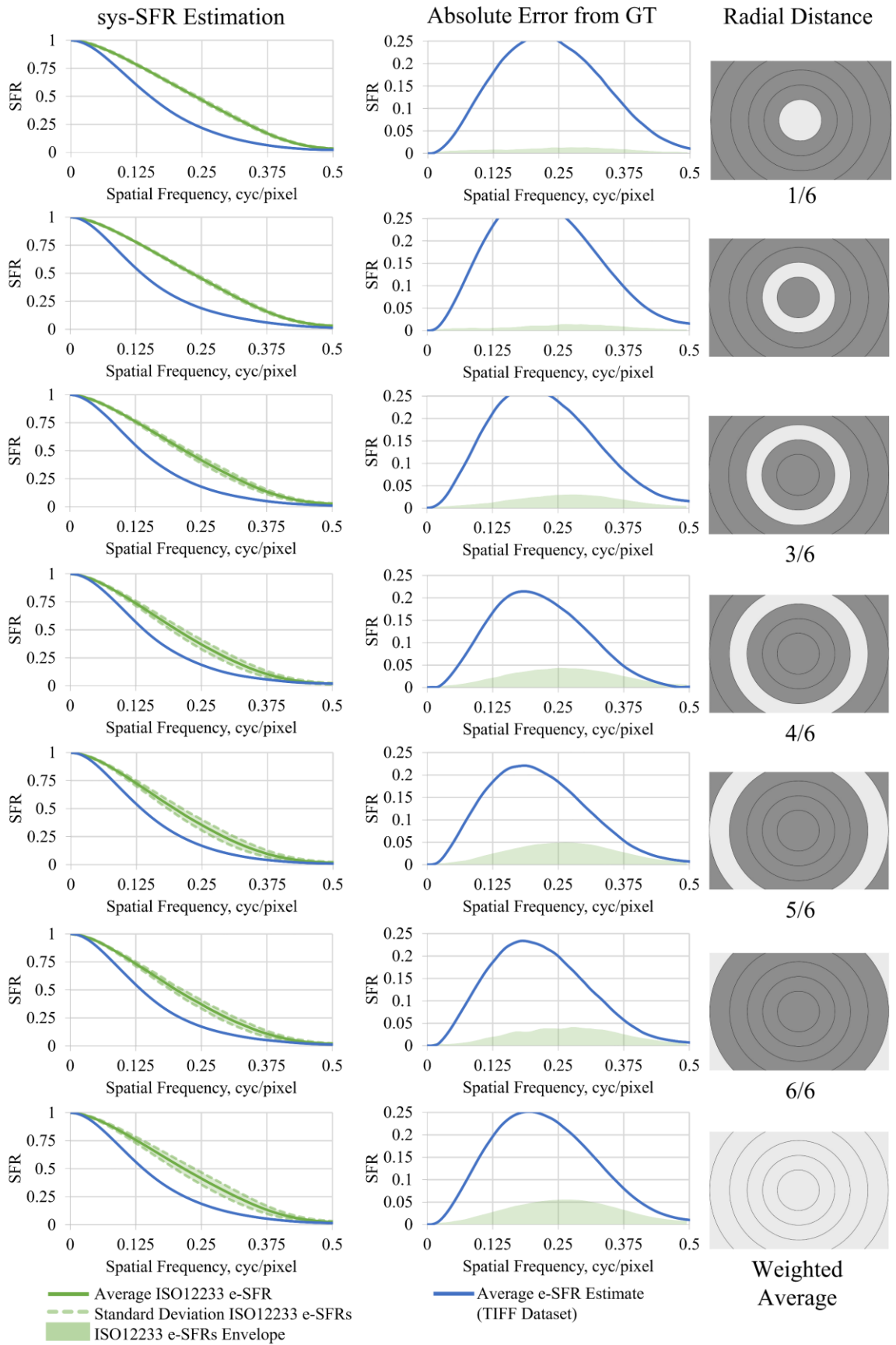


Figure C.6 DSLR 2 vertical system e-SFR estimations using the 50th percentile of the LSF FWHM distribution

C.3 Smartphone e-SFR Estimation Results

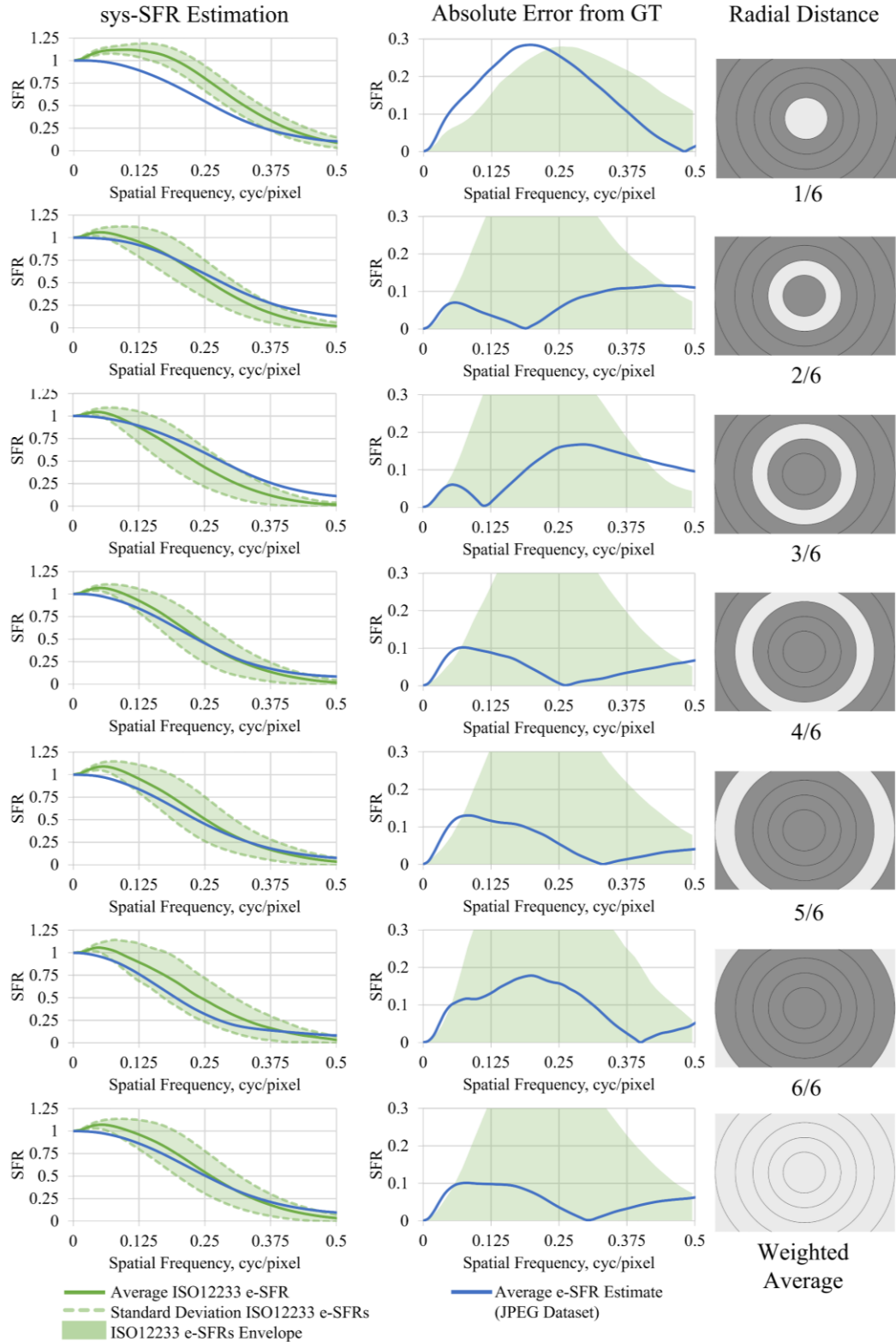


Figure C.7 Smartphone vertical system e-SFR estimations using the 10th percentile of the LSF FWHM distribution

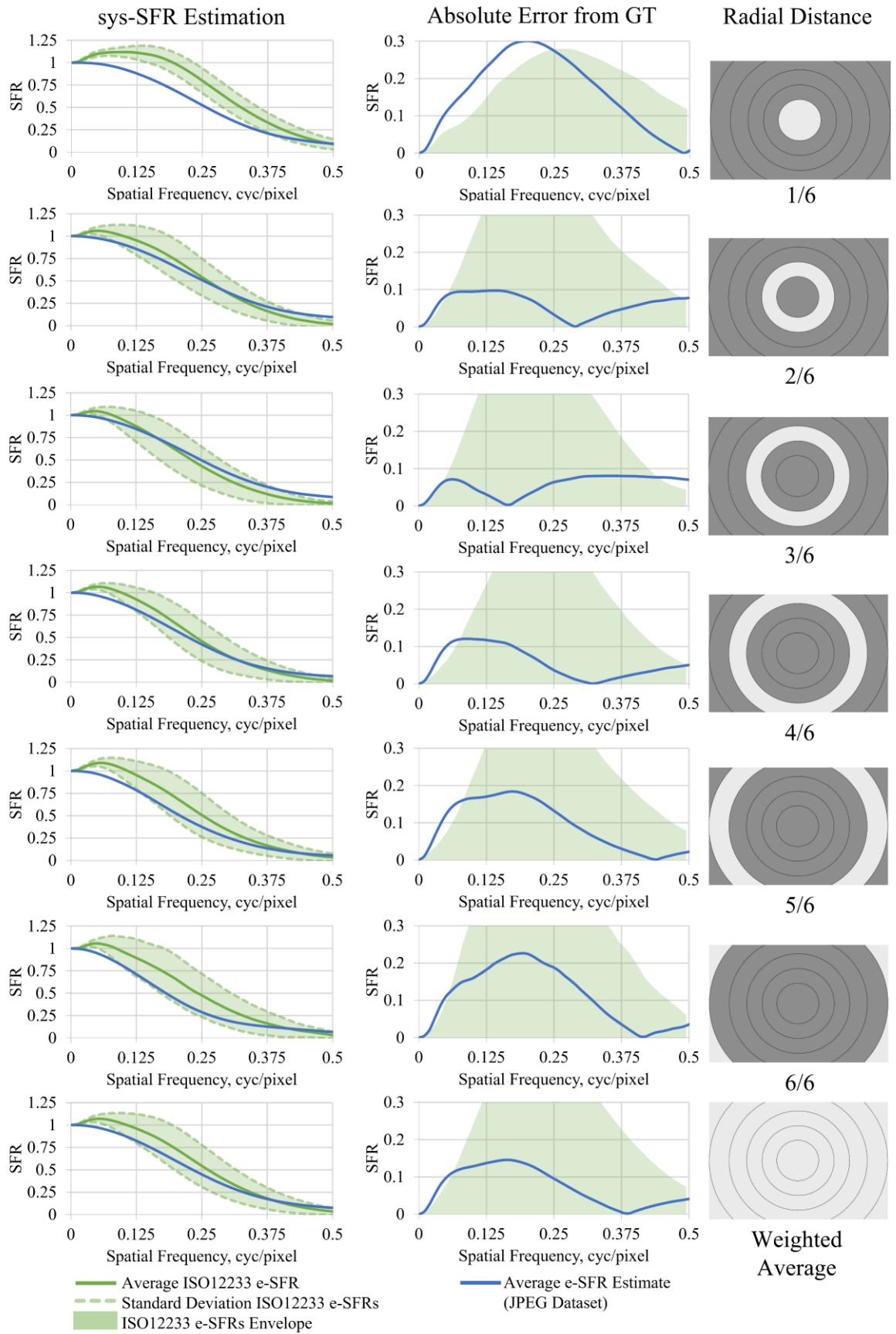


Figure C.8 Smartphone vertical system e-SFR estimations using the 20th percentile of the LSF FWHM distribution

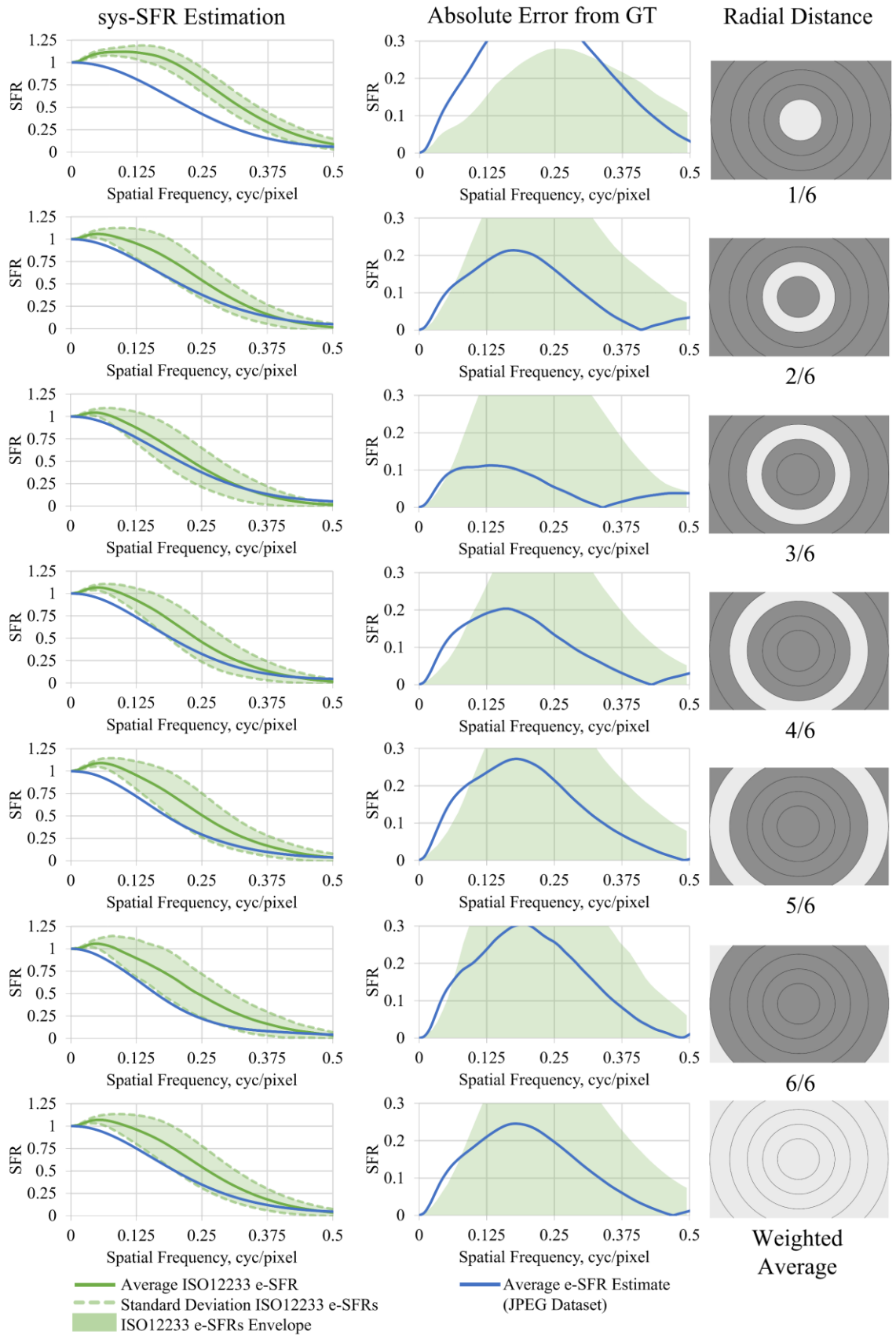


Figure C.9 Smartphone vertical system e-SFR estimations using the 50th percentile of the LSF FWHM distribution5

Appendix D Tangential DSLR e-SFR Estimation

Figures D.1 and D.2 contain plots of the tangential DSLR 1 and DSLR 2 system e-SFR estimates at 25%, 50% and 75% of the Nyquist frequency.

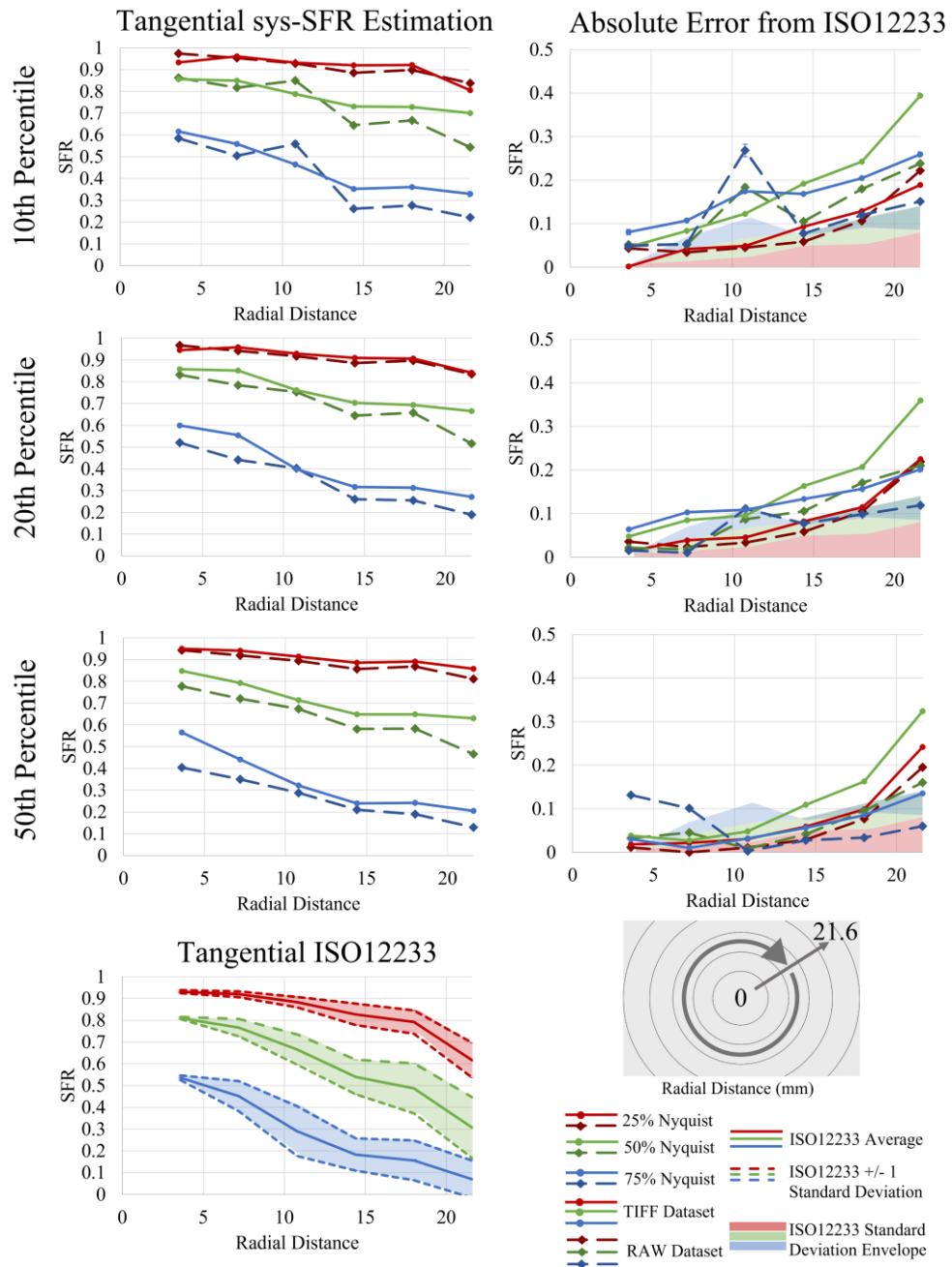


Figure D.1 25%, 50% and 75% Nyquist of the DSLR 1 tangential TIFF and RAW system e-SFR estimates are plotted against the radial distance. The estimates are derived from the 10th, 20th and 50th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO12233 e-SFR is also plotted, with the standard deviation of the ISO12233 shaded.

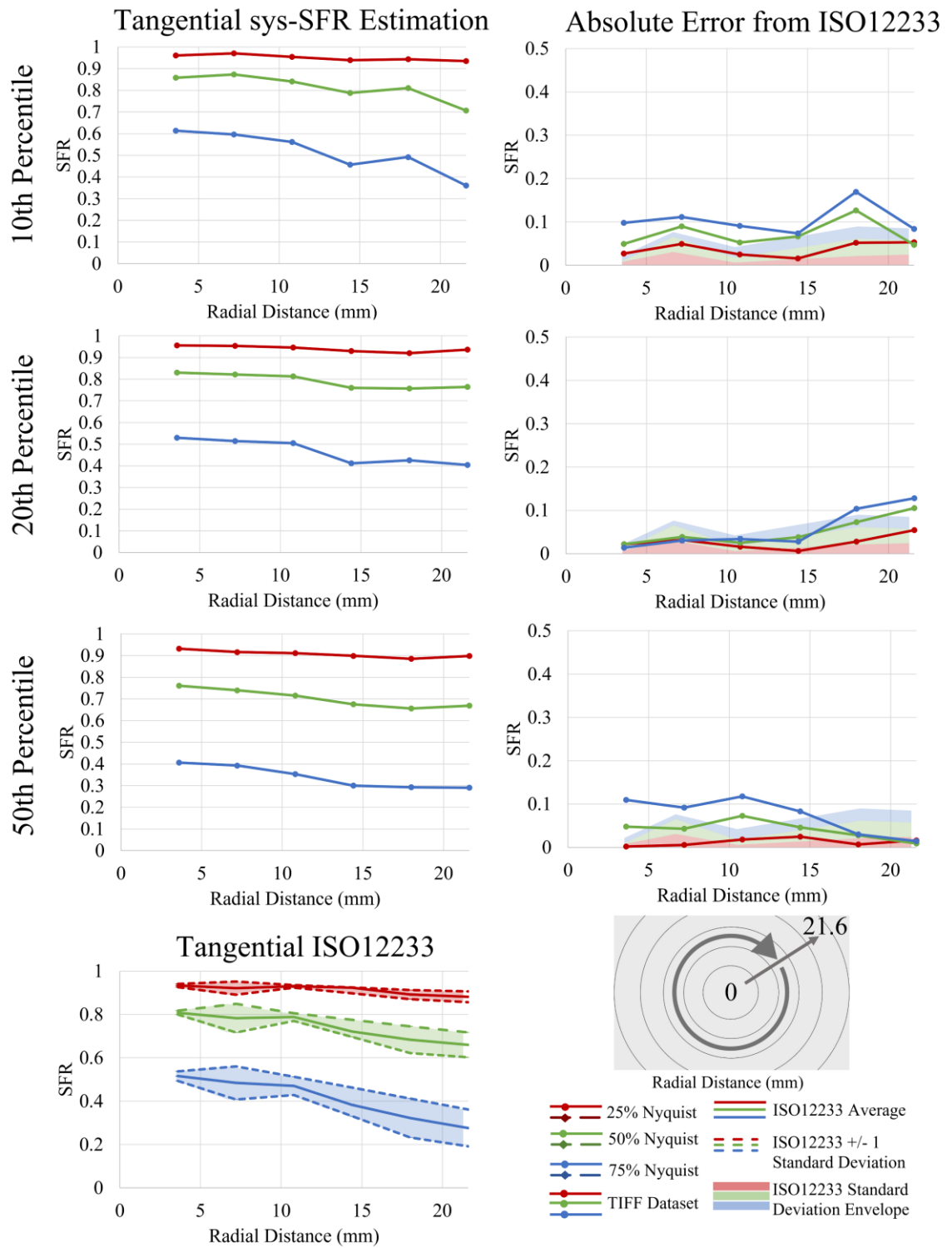


Figure D.2 25%, 50% and 75% Nyquist of the DSLR 2 tangential system e-SFR estimates are plotted against the radial distance. The estimates are derived from the 10th, 20th and 50th percentile of the LSF FWHM distribution. The absolute error of the estimated system e-SFR from the mean ISO12233 e-SFR is also plotted, with the standard deviation of the ISO12233 shaded

Appendix E MATLAB Code

The MATLAB code used in this thesis to obtain NS-SFR and subsequently estimate the system e-SFR can be found at: www.github.com/OlivervZ11/NS-SFR.

The instructions for implementing the code, the *MathWorks* Toolbox requirements, and the installation guide for the MATLAB data visitation app is found below.

E.1 Instructions & Implementation Guide

E.1.1 Requirements

MATLAB and the following MATLAB Toolboxes:

- Image Processing Toolbox
- Parallel Computing Toolbox
- Statistics and Machine Learning Toolbox

SFRMAT4

Sfrmat4 was written by P. D. Burns and is available at [87]. Throughout the provided MATLAB code, *sfrmat4.m* is used to measure the e-SFR via the slanted edge method. Minor adjustments were made to the code to output edge angle, contrast and whether clipping is present. Also, error flags are placed in the code to catch and deselect unsuitable natural scene step-edges.

E.1.2 Guide

1. Part 1 – The NS-SFR extraction:

Running *Pt1_NSSFR_Extraction.m* isolates step edges from a dataset of images. Before running the code, ensure the image dataset is stored in a folder and that all images are taken with the same camera system, lens, and aperture.

When initialising the code, you will first be prompted to state whether the dataset is either RAW or TIFF image format (see Figure E.1). The TIFF format is a standard .tif file, whilst

the RAW must be a .dng. The provided DNG image reader (*imreadDNG.m*) is based on reading a converted Nikon NEF RAW file [163]; therefore, the DNG image reader may need to be modified depending on the camera model used to capture the dataset.

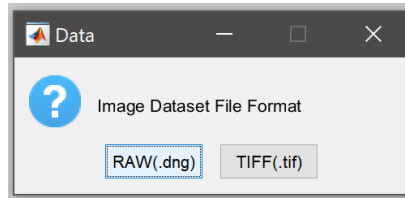


Figure E.1 Select dataset file format

After selecting a file format, you will need to select the folder that the dataset resides. Then a second directory panel will appear for you to select a folder where the result (.mat) files will be saved, a file will be generated per image, therefore it is recommended to have a dedicated results folder directory. The code will then run through processing the images and measuring the NS-SFRs.

Depending on the number of images, pixel resolution, and CPU cores available, this script may take time to run.

2. Part 2 – System e-SFR Estimation:

To estimate the system e-SFR from the extracted NS-SFR data, run the *Pt2_ESFR_Estimation.m* script. You will be prompted to select which e-SFR orientation you want to estimate, Horizontal & Vertical or Sagittal & Tangential, or both orientation groups (see Figure E.2). Note, the NS-SFRs are extracted in the Horizontal & Vertical orientation; thus, the conversion to Sagittal & Tangential will extend the computation time.

Select the directory containing the .mat files from *Pt1_NSSFR_Extraction.m*.

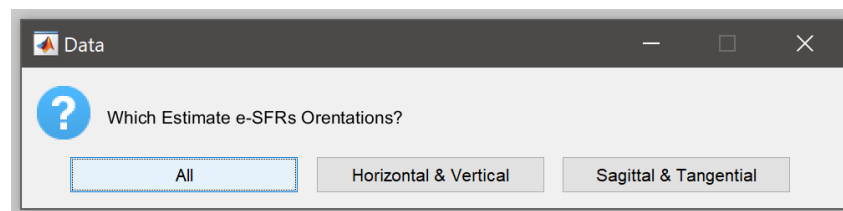


Figure E.2 Select which orientations to estimate

Depending on the orientation chosen, you will be asked to load ISO12233 Test chart data for the chosen orientations (see Figures E.3 and E.4). If ‘Yes’, choose the appropriately laid out .mat file; see Section 4 of this guide.

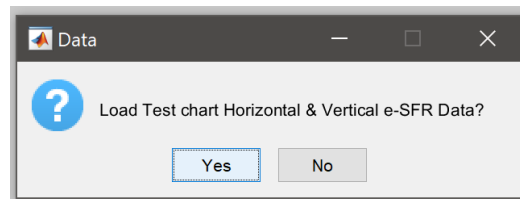


Figure E.3 Use Horizontal & Vertical test chart data?

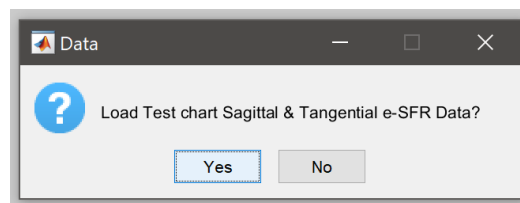


Figure E.4 Use Sagittal & Tangential test chart data?

The script will run and save a file named *eSFR_Estimation.mat* into the selected results directory.

3. Plotting the Results:

A MATLAB app user interface is provided to assist with visualising the data. To install the app in MATLAB, go to the App Tab in the top left, click Install App, and select *Plot_NSSFR_Data.mlappinstall*. The app *Plot NSSFR Data* should now be in your Apps dropdown.

Starting the app, a user interface Figure will open (see Figure E.5). Next, click the load button and select the *eSFR_Estimation.mat* file, the text under the button will inform you once loaded. A popup window will open when loading the data to ask for the camera’s pixel pitch (see Figure E.6), only required for Sagittal & Tangential plots.

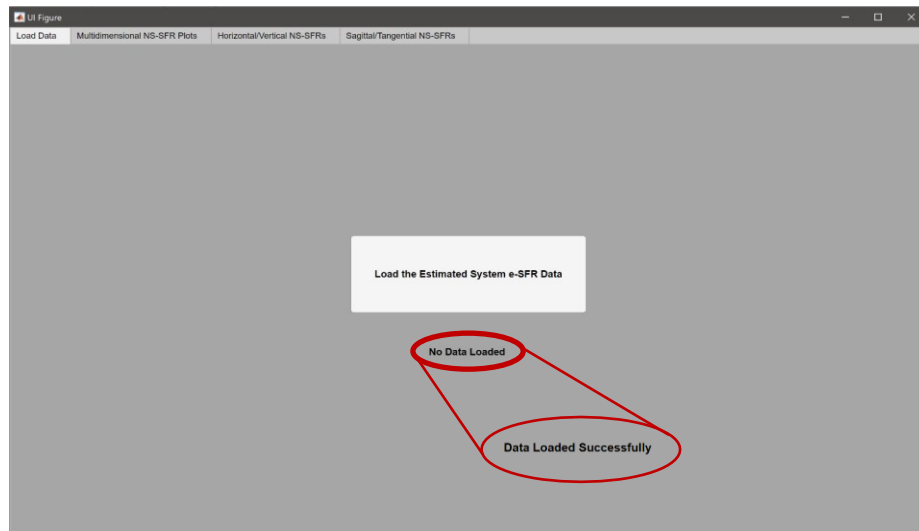


Figure E.5 Load Tab of 'Plot NSSFR Data' user interface

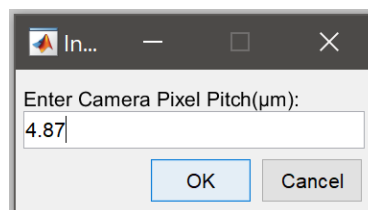


Figure E.6 Insert the camera pixel pitch

The three tabs in the user interface (see Figure E.7- E.9) provide the plots of:

- 1) the NS-SFRs binned to a multidimensional coordinate system, visualising the variation in the data,
- 2) the Horizontal & Vertical estimates,
- 3) the Sagittal & Tangential estimates.

In each tab, parameters can be adjusted, including the radial segment of the frame, the colour channel (if from RAW images), orientation and spatial frequency. In addition, if there is ISO12233 test chart data, the absolute estimation error can be displayed.

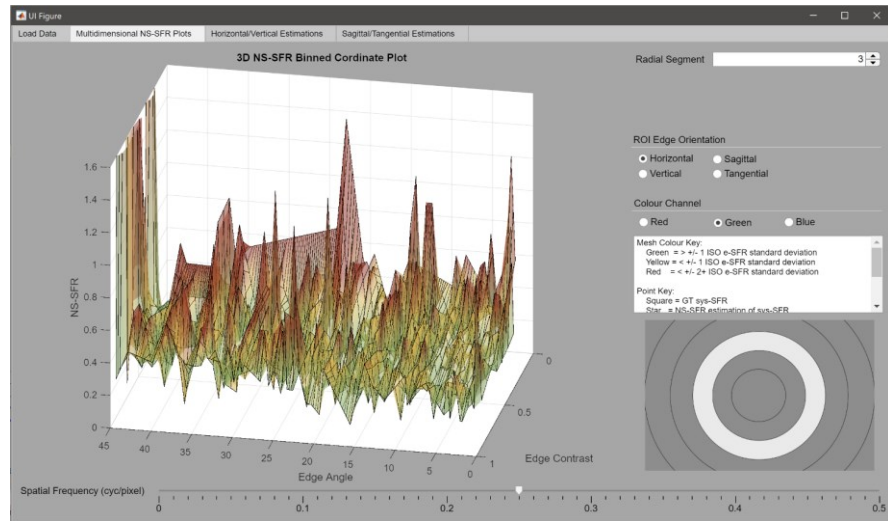


Figure E.7 Multidimensional NS-SFR Plotting Tab

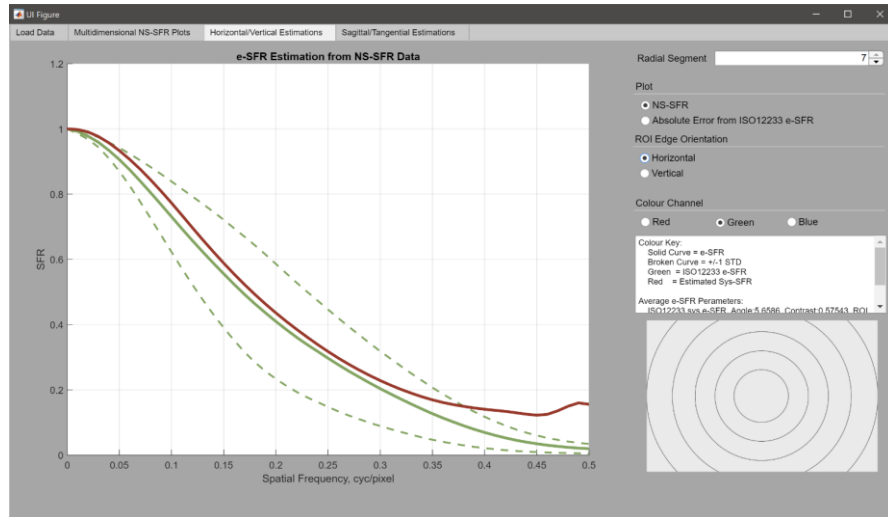


Figure E.8 Horizontal & Vertical Plotting Tab

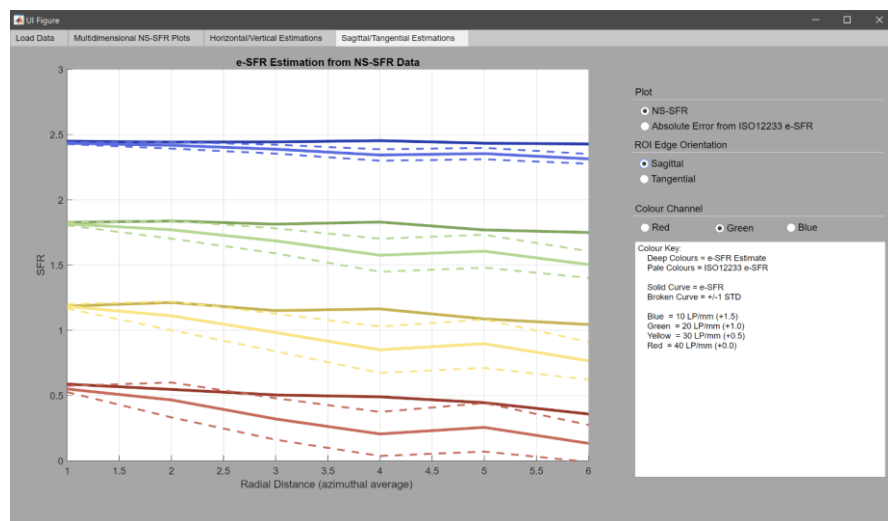


Figure E.9 Sagittal & Tangential Plotting Tab

4. Measuring the ISO12233 Test Chart e-SFR:

If test chart e-SFR data is used, the data must be correctly stored in a cell array. To achieve this, *ManualROISelection.m* is provided. Running this script will open a user interface to select step-edge from test chart images. First, store the test chart captures in a folder and select the directory when prompted. After the script loads the images from that directory, it will display the first image. Next, move the red selector to each chart Horizontal/Sagittal edges and click *Select*. Once all Horizontal/Sagittal edges are selected from the displayed image, click *Next*. If multiple test chart images are in the directory, the user interface will display the next image to repeat the Horizontal/Sagittal edge selection. After the Horizontal/Sagittal edges are selected from the images, the user interface will display the first image for you to repeat the selection process, but for the Vertical/Tangential edges.

Clicking 'Finish' will close the user interface and will ask you to save the data. The e-SFRs are measured using *sfrmat4* and are stored ready for Part 2 estimation.

The test chart should fill the frame, obtaining many e-SFRs across the field of view. Failing this, multiple captures of the test chart can be used, moving it across the field of view. Many step edges across each radial segment will provide a more accurate standard deviation measurement to assess the e-SFR estimate.

Abbreviations

1D	1-Dimensions/Dimensional
2D	2-Dimensions/Dimensional
3D	3-Dimensions/Dimensional
4D	4-Dimensions/Dimensional
5D	5-Dimensions/Dimensional
6D	6-Dimensions/Dimensional
ADC	Analogue-to-Digital Conversion
APSO	Adaptive Particle Swarm Optimization
BSI	British Standards Institute
CFA	Colour Filter Array
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CRT	Cathode Ray Tube
cyc/pixel	Cycles per pixel
DFT	Discrete Fourier Transform
DNN	Deep Neural Networks
DSLR	Digital Single Lens Reflex
ESF	Edge Spread Function
e-SFR	Edge SFR
EXIF	Exchangeable Image File Format
FPN	Fixed Pattern Noise
FWHM	Full Width at Half Maximum
GPU	Graphics Processing Unit
HVS	Human Visual System
IMU	Inertial Measurement Unit
IQ	Image Quality
IQM	Image Quality Metric
IR	Infrared
ISO	International Organization for Standardization
ISP	Image Signal Processor
LCD	Liquid Crystal Displays
LSF	Line Spread Function
LUT	Look-up-Table
MAE	Mean Absolute Error
MTF	Modulation Transfer Function
NaN	Not a Number
NSS	Natural Scene Statistics
NS-SFR	Natural Scene derived SFR
OEFC	Opto-Electric Conversion Function

OTF	Optical Transfer Function
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PPI	Pixels per inch
PRNU	Pixel Response Non-Uniformity
PSF	Point Spread Function
PTF	Phase Transfer Function
QE	Quantum Efficiency
RGB	Red Green Blue
RMS	Root-Mean-Square
ROI	Region of Interest
SE	Standard Error
SFR	Spatial Frequency Response
SNR	Signal-to-Noise Ratio
SPD	Signal and Processing Dependant
sRGB	Standard RGB

References

- [1] British Standard Institute, “PD ISO/TS 23564:2020 Image technology colour management — Evaluating colour transform accuracy in ICC profiles,” *BSI Standards Publication*, pp. 1–16, 2020.
- [2] British Standard Institute, “BS ISO 14524:2009 Photography - Electronic still-picture cameras - Methods for measuring opto-electronic conversion functions (OECFs),” *BSI Standards Publication*, pp. 1–23, 2009.
- [3] British Standard Institute, “BS ISO 17850:2015 Photography — Digital cameras — Geometric distortion (GD) measurements,” *BSI Standards Publication*, p. 60, 2015.
- [4] British Standard Institute, “BS ISO 15739:2017 Photography — Electronic still-picture imaging — Noise measurements,” *BSI Standards Publication*, p. 42, 2017.
- [5] British Standard Institute, “BS ISO 12233:2017 Photography - Electronic still picture imaging - Resolution and spatial frequency responses,” *BSI Standards Publication*, pp. 1–62, 2017.
- [6] R. Jenkin, “Noise, sharpness, resolution and information,” in *The Manual of Photography*, Tenth Edit., E. Allen and S. Triantaphillidou, Eds. New York & London: Focal Press, Taylor & Francis Group, 2009, pp. 433–456.
- [7] R. Jenkin, “Images and Image Formation,” in *The Manual of Photography*, 10th ed., Burlington: Focal Press, 2011, pp. 119–138.
- [8] A. Ignatov *et al.*, “Learned smartphone ISP on mobile NPUs with deep learning, mobile AI 2021 challenge: Report,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 2503–2514, 2021, doi: 10.1109/CVPRW53098.2021.00284.
- [9] M. Delbracio, D. Kelly, M. S. Brown, and P. Milanfar, “Mobile Computational Photography: A Tour,” *Annual Review of Vision Science*, vol. 7, pp. 571–604, 2021, doi: 10.1146/annurev-vision-093019-115521.
- [10] J. C. Dainty and R. Shaw, “The Modulation Transfer Function,” in *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*, Academic Press, 1974, pp. 232–275.

-
- [11] International Imaging Industry Association, “Camera Phone Image Quality – Phase 2 Introduction,” *International Imaging Industry Association*, p. 3, 2009.
- [12] R. Branca, S. Triantaphillidou, and P. Burns, “Texture MTF from images of natural scenes,” *IS&T Image Quality and System Performance XIV*, pp. 113–120, 2017.
- [13] 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,” *Journal of Imaging Science and Technology*, 2020.
- [14] E. W. S. Fry, “Scene-Dependency of Spatial Image Quality Metrics,” University of Westminster, 2020.
- [15] M. Bauer, V. Volchkov, M. Hirsch, and B. Schölkopf, “Automatic Estimation of Modulation Transfer Functions,” in *IEEE International Conference on Computational Photography (ICCP)*, 2018, pp. 1–40.
- [16] M. Bohra, S. Maheshwari, and V. Gandhi, “TextureToMTF: predicting spatial frequency response in the wild,” *Signal, Image and Video Processing*, vol. 14, no. 6, pp. 1163–1170, 2020, doi: 10.1007/s11760-020-01656-w.
- [17] 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 Image Quality and System Performance XVII*, no. 9, pp. 1–32, 2020, doi: 10.2352/J.ImagingSci.Technol.2019.63.6.060407.
- [18] J. C. Dainty and R. Shaw, “Fourier Transforms, and the Analysis of Image Resolution and Noise,” in *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*, London: Academic Press, 1974, pp. 190–321.
- [19] E. Allen, *Spatial Image Processing*, Tenth Edit. Elizabeth Allen & Sophie Triantaphillidou, 2011.
- [20] R. B. Jenkin, “On the Application of the Modulation Transfer Function to Discrete Imaging Systems,” University of Westminster, 2001.
- [21] N. Koren, “Log Frequency - Imatest Documentation - Current v2021.1,” *Imatest*, 2021. https://www.imatest.com/docs/log_f/ (accessed Aug. 21, 2021).
- [22] S. E. Reichenbach, S. K. Park, and R. Narayanswamy, “Characterizing digital image acquisition devices,” *Optical Engineering*, vol. 30, no. 2, p. 170, 1991, doi: 10.1117/12.55783.

-
- [23] W. F. Hsu, Y. C. Hsu, and K. W. Chuang, "Measurement of the spatial frequency response (SFR) of digital still-picture cameras using a modified slanted edge method," *Proceedings of SPIE*, pp. 96–103, 2000.
- [24] A. M. Ford, "Relationships Between Image Quality and Image Compression," University of Westminster, 1997.
- [25] F. Cao, F. Guichard, and H. Hornung, "Measuring Texture Sharpness of A Digital Camera," 2009, doi: 10.1117/12.805853.
- [26] F. Cao, F. Guichard, and H. Hornung, "Dead leaves model for measuring texture quality on a digital camera," 2010, doi: 10.1117/12.838902.
- [27] P. D. Burns, "Refined Measurement of Digital Image Texture Loss," in *SPIE Image Quality and System Performance X*, 2013, vol. 8653, doi: 10.1117/12.2004432.
- [28] U. Artmann, "Image quality assessment using the dead leaves target : experience with the latest approach and further investigations," in *SPIE Digital Photography XI*, 2015, vol. 9404, pp. 1–15, doi: 10.1117/12.2079609.
- [29] 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," vol. 7537, p. 75370D, 2010, doi: 10.1117/12.838698.
- [30] Image Engineering, "TE265 Dead-Leaves Test Chart," 2021. <https://www.image-engineering.de/products/charts/all/582-te265> (accessed Aug. 22, 2021).
- [31] Image Engineering, "TE276 Dead-Leaves Test Chart," 2021. <https://www.image-engineering.de/products/charts/all/401-te276> (accessed Aug. 22, 2021).
- [32] 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).
- [33] L. Kirk, P. Herzer, U. Artmann, and D. Kunz, "Description of texture loss using the dead leaves target: current issues and a new intrinsic approach," *Digital Photography X*, 2014, doi: 10.1117/12.2039689.
- [34] R. C. Sumner, R. Burada, I. Llc, and N. Kram, "The Effects of misregistration on the dead leaves crosscorrelation texture blur analysis," *Electronic Imaging, Image Quality and System Performance XIV*, pp. 121–129, 2017.

-
- [35] S. Triantaphillidou, R. E. Jacobson, and R. Fagard-Jenkin, "An Evaluation of MTF Determination Methods for 35mm Film Scanners," in *IS&T PICS Conference, Image Quality*, 1999, no. 3, pp. 231–235.
- [36] J. C. Feltz and M. A. Karim, "Modulation transfer function of charge-coupled devices," *Applied Optics*, vol. 29, no. 5, pp. 717–722, 1990, doi: 10.1364/AO.29.000717.
- [37] R. B. Fagard-Jenkin, R. E. Jacobson, and N. R. Axford, "A Novel Approach to the Derivation of Expressions for Geometrical MTF in Sampled Systems," in *Society for Imaging Science and Technology: Image Processing, Image Quality, Image Capture, Systems Conference*, 1999, pp. 225–230.
- [38] P. D. Burns, "Estimation Error in Image Quality Measurements," *Proc. SPIE, Image Quality and System Performance VIII*, vol. 7867, 2011.
- [39] E. C. Yeadon, R. A. Jones, and J. T. Kelly, "Confidence Limits for Individual Modulation Transfer Function Measurements based on the Phase Transfer Function," *Photographic Science and Engineering*, vol. 14, no. 2, pp. 153–156, 1970.
- [40] P. D. Burns, "Image Quality Concepts," in *Handbook of Digital Imaging*, M. Kriss, Ed. John Wiley & Sons, Inc, 2015, pp. 1–47.
- [41] J. C. Dainty and R. Shaw, "Input/Output Relationships for Conventional Photographic Processes: Experimental Observables," in *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*, 1974, pp. 33–67.
- [42] R. C. Gonzalez and R. E. Woods, "Sampling in the Frequency Domain," in *Digital Image Processing*, 3rd ed., Upper Saddle River: N.J: Prentice Hall, 2008, pp. 199–311.
- [43] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, New Jersey: Prentice Hall, 1989.
- [44] British Standard Institute, "BS ISO 12231:2012 Photography — Electronic still picture imaging — Vocabulary," *BSI Standards Publication*, 2012.
- [45] 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.
- [46] C. N. Proudfoot, *Handbook of Photographic Science and Engineering*, 2nd ed. IS&T, 1997.
- [47] J. E. Farrell, P. B. Catrysse, and B. A. Wandell, "Digital camera simulation," *Applied Optics*, vol. 51, p. 43, doi: 10.1364/AO.51.000A80.

-
- [48] J. Chen *et al.*, “Digital camera imaging system simulation,” *IEEE Transactions on Electron Devices*, vol. 56, no. 11, pp. 2496–2505, 2009, doi: 10.1109/TED.2009.2030995.
- [49] Imatest, “Using SFR Part 1 (v2021.2),” 2021. https://www.imatest.com/docs/sfr_instructions/ (accessed Jan. 07, 2022).
- [50] S. Triantaphillidou, *Tone reproduction*, Tenth Edit., vol. 4. Elizabeth Allen & Sophie Triantaphillidou, 2011.
- [51] M. Langford and E. Bilissi, “Digital Imaging Systems,” in *Langford’s Advanced Photography: The Guide for Aspiring Photographers*, 8th ed., Amsterdam ; Boston: Elsevier: Focal Press, 2011, pp. 170–196.
- [52] Y. Okano, “Influence of Image Enhancement Processing on SFR of Digital Cameras,” *Opt. Rev.*, vol. 5, no. 6, pp. 358–362, 1998.
- [53] P. D. Burns, “Tone-transfer (OECF) characteristics and spatial frequency response measurements for digital cameras and scanners,” in *SPIE-IS&T Electronic Imaging Symposium*, 2005, vol. 5668, pp. 123–128, doi: 10.1117/12.585884.
- [54] Adobe Systems Inc., “Adobe RGB (1998) Color Image Encoding,” *Adobe Systems*, pp. 1–20, 2005, [Online]. Available: <papers://52bf690d-fd16-48df-a6bf-d4ec950402d8/Paper/p1453>.
- [55] International Electrotechnical Commission, “International Standard IEC 61966-2-1:1999 - Colour management – Default RGB colour space - sRGB,” *International Electrotechnical Commission*, p. 13, 1999.
- [56] P. D. Burns, “Slanted-Edge MTF for Digital Camera and Scanner Analysis,” *Proc. IS&T 2000 PICS Conference*, vol. 3, pp. 135–138, 2000, doi: 10.1.1.149.2643.
- [57] U. Artmann, “Linearization and Normalization in Spatial Frequency Response Measurement,” *Electronic Imaging, Image Quality and System Performance XIII*, vol. 8, pp. 1–6, 2016, doi: doi.org/10.2352/ISSN.2470-1173.2016.13.IQSP-011.
- [58] P. D. Burns and D. Williams, “Sampling Efficiency in Digital Camera Performance Standards,” *Proc. SPIE IS&T*, vol. 6808, pp. 1–5, 2008, doi: 10.1117/12.766359.
- [59] R. B. Jenkin, R. E. Jacobson, M. A. Richardson, and I. C. Luckraft, “Analytical MTF Bounds and Estimate for SFR in Discrete Imaging Array Due to Non-Stationary Effects,” *Journal of Imaging Science and Technology*, vol. 47, no. 3, pp. 200–208, 2003.

-
- [60] R. B. Jenkin, R. E. Jacobson, M. A. Richardson, and I. C. Luckraft, "Extension of the ISO 12233 SFR measurement technique to provide MTF bounds for critical imaging arrays," in *IS&T/SPIE Electronic Imaging*, 2005, p. 10, doi: 10.1117/12.593081.
- [61] J. K. M. Roland, "A study of slanted-edge MTF stability and repeatability," *SPIE Proceedings/IS&T Image Quality and System PerformanceXII*, vol. 9396, pp. 1–9, 2015, doi: 10.1117/12.2077755.
- [62] Y. Okano, "Old Influence of Non-Linear Image Processing on Spatial Frequency Response of Digital Still Cameras," *Opt. Rev.*, vol. 5, no. 6, pp. 358–362, 1998, doi: 10.1007/s10043-998-0358-4.
- [63] D. Williams, "Benchmarking of the Slanted-edge Spatial Frequency Response Plug-in," in *IS&T's PICS Conference*, 1998, pp. 133–136.
- [64] S. Birchfield, "Reverse-Projection Method for Measuring Camera MTF," *IS&T, Image Quality and System PerformanceXIV*, pp. 105–112, 2017.
- [65] K. Masaoka, "Edge-based modulation transfer function measurement method using a variable oversampling ratio," *Optics Express*, vol. 29, no. 23, pp. 37628–37638, 2021, doi: 10.1364/oe.435981.
- [66] K. Masaoka, "Real-time modulation transfer function measurement system," *SPIE Ultra-High-Definition Imaging Syst. II*, vol. 10943, no. 1094309, pp. 1–9, 2019, doi: 10.1117/12.2517201.
- [67] F. van den Bergh, "Robust edge spread function construction methods to counter poor sample spacing uniformity in the slanted-edge method," *Optical Society of America*, vol. 36, no. 7, pp. 1126–1136, 2019, doi: 10.1364/josaa.36.001126.
- [68] Imatest, "Using SFR Part 2: Running Imatest SFR (v.2021.1)," 2021. https://www.imatest.com/docs/sfr_instructions2/ (accessed Dec. 04, 2020).
- [69] P. D. Burns and D. Williams, "Refined slanted-edge measurement for practical camera and scanner testing," *IS&T PICS Conference*, pp. 191–195, 2002.
- [70] J. N. A. Laitre, "Determination of the chemical spread function of Kodak High Contrast Copy 5069 Emulsion," Rochester Institute of Technology, 1979.
- [71] IEEE Standards Association, "IEEE 1858: Standard for Camera Phone Image Quality (CPIQ)." pp. 1–146, 2016.

-
- [72] R. Jenkin, R. E. Jacobson, and N. R. Axford, "Partial Determination of Non-Linear MTF via the Volterra and Wiener Theories of Non-Linear Systems," in *Proceedings of IS&T PICS 2001 Conference*, 2001, pp. 233–238.
- [73] S. Ray, "Aberrations - defects in imaging systems," in *Applied Photographic Optics*, 3rd ed., Focal Press, 2002, pp. 82–98.
- [74] Imatest, "Using eSFR ISO Part 2 (v.2021.1)," 2021. https://www.imatest.com/docs/esfriso_instructions2/ (accessed Dec. 07, 2020).
- [75] BSI Standards Publication, "BS ISO 20462-3:2012 Psychophysical experimental methods for estimating image quality Part 3: Quality ruler method," *BSI Standards Publication*, pp. 1–23, 2012.
- [76] O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, "Camera System Performance Derived from Natural Scenes," in *IS&T Electronic Imaging: Image Quality and System Performance XVII*, 2020, pp. 241-1-241–10, [Online]. Available: <https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-241>.
- [77] Imatest, "Correcting non-uniformity in slanted-edge MTF measurements," *Imatest*, 2020. <https://www.imatest.com/2018/08/edge-mtf-nonuniformity/> (accessed Mar. 10, 2021).
- [78] 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.
- [79] N. Koren, R. C. Sumner, and H. Koren, "Measuring MTF with wedges: Pitfalls and best practices," 2017, doi: 10.2352/ISSN.2470-1173.2017.19.AVM-451.
- [80] N. L. Koren, "Correcting Misleading Image Quality Measurements," 2020, doi: 10.2352/ISSN.2470-1173.2020.9.IQSP-242.
- [81] Imatest, "Using SFRplus, Part 3: Results (v.2021.1)," 2021. https://www.imatest.com/docs/sfrplus_instructions3/ (accessed Aug. 28, 2021).
- [82] H. H. Nasse, "How to Read MTF Curves," *Carl Zeiss*, pp. 1–33, 2008.
- [83] 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.

-
- [84] K. Masaoka, T. Yamashita, Y. Nishida, and M. Sugawara, "Modified slanted-edge method and multidirectional modulation transfer function estimation," *Optics Express*, vol. 22, no. 5, p. 6040, 2014, doi: 10.1364/OE.22.006040.
- [85] R. L. Baer, "The Circular Edge Spatial Frequency Response Test," *SPIE 5294, Image Quality and System Performance*, 2003, doi: 10.1117/12.524829.
- [86] P. D. Burns and D. Williams, "Camera Resolution and Distortion: Advanced Edge Fitting," *Society for Imaging Science and Technology*, vol. 15, pp. 1–5, 2018.
- [87] P. D. Burns, "sfrmat4 source code," *Burns Digital Imaging*, 2020. <http://burnsdigitalimaging.com/software/sfrmat/> (accessed Oct. 04, 2020).
- [88] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283–2302, 2007, doi: 10.1016/j.sigpro.2007.04.004.
- [89] 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 Mtg.*, pp. 1–8, 2005.
- [90] S. Kabir, L. Leigh, and D. Helder, "Vicarious methodologies to assess and improve the quality of the optical remote sensing images: A critical review," *Remote Sensing*, vol. 12, no. 24, pp. 1–40, 2020, doi: 10.3390/rs12244029.
- [91] T. Choi, X. Xiong, and Z. Wang, "On-orbit lunar modulation transfer function measurements for the moderate resolution imaging spectroradiometer," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 270–277, 2014, doi: 10.1109/TGRS.2013.2238545.
- [92] K. Kohm, L. H. Parkway, and S. Louis, "Modulation Transfer Function Measurement Method and Results for the Orbview-3 High Resolution Imaging Satellite," pp. 1–6, 2003.
- [93] F. Viallefont-Robinet and D. Léger, "Improvement of the edge method for on-orbit MTF measurement," *Optics Express*, vol. 18, no. 4, pp. 3531–3545, 2010, doi: 10.1364/oe.18.003531.
- [94] 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 Conference Proceedings*, no. 1, pp. 43–47, 2013.
- [95] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

-
- [96] 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 International Symposium on Electronic Imaging: Image Quality and System Performance XV*, p. 6, 2018.
- [97] S. Haykin, "Blind equalization formulated as a self-organized learning process," *Conference Record of the Twenty-Sixth Asilomar Conference on Signals, Systems & Computers*, vol. 1, pp. 346–350, 1992, doi: 10.1109/ACSSC.1992.269177.
- [98] S. Haykin, "Blind Deconvolution," in *Adaptive Filter Theory*, 2nd ed., vol. 4, no. 11, United Kingdom, Prentice Hall, 1994, pp. 722–744.
- [99] D. Hatzinakos, "Blind Deconvolution Channel Identification and Equalization," *Control and dynamic systems*, vol. 68, pp. 279–331, Jan. 1995, doi: 10.1016/S0090-5267(06)80043-5.
- [100] A. K. Katsaggelos, S. D. Babacan, and T. Chun-Jen, "Iterative Image Restoration," *The Essential Guide to Image Processing*, pp. 349–383, Jan. 2009, doi: 10.1016/B978-0-12-374457-9.00015-9.
- [101] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution.," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995, doi: 10.1162/neco.1995.7.6.1129.
- [102] J. M. Blackledge, "Image Restoration and Reconstruction," *Digital Image Processing*, pp. 404–438, Jan. 2005, doi: 10.1533/9780857099464.3.403.
- [103] E. Y. Lam and J. W. Goodman, "Iterative statistical approach to blind image deconvolution," *Journal of the Optical Society of America*, vol. 17, no. 7, pp. 1177–1184, 2000, doi: 10.1364/josaa.17.001177.
- [104] M. Q. Pham, B. Oudompheng, J. I. Mars, and B. Nicolas, "A Noise-Robust Method with Smoothed ℓ_1/ℓ_2 Regularization for Sparse Moving-Source Mapping," *Signal Processing*, vol. 135, pp. 96–106, Jun. 2017, doi: 10.1016/J.SIGPRO.2016.12.022.
- [105] X. Ge, J. Tan, L. Zhang, and Y. Qian, "Blind image deconvolution via salient edge selection and mean curvature regularization," *Signal Processing*, vol. 190, 2022, doi: 10.1016/j.sigpro.2021.108336.
- [106] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Revisiting blind motion deblurring using conditional adversarial networks," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 168, pp. 8183–8192, 2018, doi: 10.1016/j.sigpro.2019.107338.

-
- [107] J. Kotera, F. Sroubek, and V. Smidl, “Improving Neural Blind Deconvolution,” *IEEE International Conference on Image Processing (ICIP)*, no. 2, pp. 1954–1958, 2021, doi: 10.1109/icip42928.2021.9506502.
- [108] Y. Zheng, W. Huang, Y. Pan, and M. Xu, “Optimal PSF estimation for simple optical system using a wide-band sensor based on psf measurement,” *Sensors*, no. 10, 2018, doi: 10.3390/s18103552.
- [109] M. Bohra, S. Maheshwari, and V. Gandhi, “Generalized SFR Dataset (GSFR),” *GitHub*, 2020. <https://github.com/MurtuzaBohra/TextureToMTF> (accessed Nov. 22, 2021).
- [110] S. Ray, “Stereoscopic (3D) Photography,” in *Scientific Photography and Applied Imaging*, Focal Press, 1999, pp. 248–270.
- [111] C. C. Su, L. K. Cormack, and A. C. Bovik, “Depth estimation from monocular color images using natural scene statistics models,” *IEEE 11th IVMSP Workshop: 3D Image/Video Technologies and Applications, IVMSP*, pp. 1–4, 2013, doi: 10.1109/IVMSPW.2013.6611900.
- [112] L. He, Q. Dong, and G. Wang, “Fast depth extraction from a single image,” *International Journal of Advanced Robotic Systems*, vol. 13, no. 6, pp. 1–11, 2016, doi: 10.1177/1729881416663370.
- [113] J. H. Lee and C. S. Kim, “Monocular depth estimation using relative depth maps,” in *Proceedings of the IEEE Computer Society Conference*, 2019, pp. 9721–9730, doi: 10.1109/CVPR.2019.00996.
- [114] I. Alhashim and P. Wonka, “High Quality Monocular Depth Estimation via Transfer Learning,” 2018, [Online]. Available: <http://arxiv.org/abs/1812.11941>.
- [115] L. Lin, G. Huang, Y. Chen, L. Zhang, and B. He, “Efficient and high-quality monocular depth estimation via gated multi-scale network,” *IEEE Access*, vol. 8, pp. 7709–7718, 2020, doi: 10.1109/ACCESS.2020.2964733.
- [116] C. A. Shah and W. Schickler, “Automated Blur Detection and Removal in Airborne Imaging Systems Using IMU Data,” in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012, pp. 321–323, doi: 10.5194/isprsarchives-xxxix-b1-321-2012.
- [117] M. Langford and E. Bilissi, “Image Sensors,” in *The Guide for Aspiring Photographers*, 8th ed., Amsterdam ; Boston: Elsevier: Focal Press, 2011, pp. 143–169.

-
- [118] G. C. Holst, "Array Performance," in *CCD Arrays, Cameras, and Displays*, 2nd ed., Bellingham, Wash., USA: SPIE Optical Engineering, 1998, pp. 102–145.
- [119] R. C. Gonzalez and R. E. Woods, "Colour Image Processing," in *Digital Image Processing*, 3rd ed., Upper Saddle River, N.J: Prentice Hall, 2008, pp. 394–460.
- [120] R. Jenkin, "Speed and Sensitivity," in *The Manual of Photography*, 10th ed., Burlington: Focal Press, 2011, pp. 365–375.
- [121] J. E. Farrell and B. A. Wandell, "Image Systems Simulation," *SID Symposium Digest of Technical Papers*, vol. 46, no. 1, pp. 1–28, 2015, doi: 10.1002/sdtp.10297.
- [122] J. C. Dainty and R. Shaw, "Detective Quantum Efficiency, Signal-to-Noise Ratio, and the Noise-Equivalent Number of Quanta," in *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*, Academic Press, 1974, pp. 152–189.
- [123] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8 (6), pp. 679–698, 1986, doi: 10.1109/TPAMI.1986.4767851.
- [124] 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 Imaging*, vol. 65, no. 5, pp. 60402-1-60402–16, 2022.
- [125] L. Ding and A. Goshtasby, "On the canny edge detector," *Pattern Recognition*, vol. 34, no. 3, pp. 721–725, 2001, doi: 10.1016/S0031-3203(00)00023-6.
- [126] R. M. Haralick and L. G. Shapiro, "Computer and Robot Vision," Volume II., Addison-Wesley, 1992, pp. 316–317.
- [127] J. P. Lewis, "Fast Normalized Cross-Correlation," *Vision Interface*, vol. 1995, no. 1, pp. 1–7, 1995.
- [128] M. Fang, G. Yue, and Q. Yu, "The Study on An Application of Otsu Method in Canny Operator," *International Symposium on Information Processing*, vol. 2, no. 4, pp. 109–112, 2009.
- [129] Y. Huo, G. Wei, Y.-D. Zhang, and L. Wu, "An Adaptive Threshold for the Canny Operator of Edge Detection," *International Conference on Image Analysis and Signal Processing*, pp. 109–112, 2010, doi: 10.1109/IASP.2010.5476095.

-
- [130] R. K. Sidhu, "Improved Canny Edge Detector in Various Color Spaces," *International Conference on Reliability, Infocom Technologies and Optimization*, vol. 3, pp. 1–6, 2014.
- [131] Rachmawan, "Canny Edge Detection," *MATLAB Central File Exchange*, 2014. <https://uk.mathworks.com/matlabcentral/fileexchange/46859-canny-edge-detection> (accessed Jun. 03, 2019).
- [132] MATLAB, "normxcorr2," *MathWorks*, 2006. <https://uk.mathworks.com/help/images/ref/normxcorr2.html> (accessed Jul. 29, 2021).
- [133] MATLAB, "regionprops," *MathWorks*, 2006. <https://uk.mathworks.com/help/images/ref/regionprops.html> (accessed Jul. 28, 2021).
- [134] P. D. Burns and D. Williams, "Using Slanted Edge Analysis for Color Registration Measurement," *Society for Imaging Science and Technology: Image Processing, Image Quality, Image Capture, Systems Conference*, pp. 51–53, 1999, doi: 10.1/1.89.5370.
- [135] 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 International Symposium on Electronic Imaging 2020 Image Quality and System Performance*, pp. 240–246, 2020, doi: <https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-240>.
- [136] International Telecommunication Union (ITU), "Recommendation BT.601-7: Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios," 2011, [Online]. Available: <http://www.itu.int/ITU-R/go/patents/en>.
- [137] MathWorks, "Convert RGB image or colormap to grayscale - MATLAB rgb2gray," *MathWorks*, 2021. <https://uk.mathworks.com/help/matlab/ref/rgb2gray.html> (accessed Dec. 12, 2021).
- [138] MathWorks, "2-D and 3-D grids - MATLAB meshgrid," *MathWorks*, 2021. <https://uk.mathworks.com/help/matlab/ref/meshgrid.html> (accessed Dec. 17, 2021).
- [139] Phantom, "Boosting Sensitivity," *Phantom High Speed*, 2016. <https://phantomhighspeed-service.force.com/servlet/servlet.FileDownload?file=00P1N00000dSuGDUA0&oid=00D1N000002Br2EUAS> (accessed Dec. 17, 2021).
- [140] Y. Li, J. Lu, L. Wang, and T. Yahagi, "Noise Removal for Degraded Images with Poisson Noise Using M-Transformation and BayesShrink Method," *Electron. Commun. Japan, Part III Fundam. Electron. Sci. (English Transl. Denshi Tsushin Gakkai Ronbunshi)*, vol. 90, no. 11, pp. 503–512, 2007, doi: 10.1002/ecjc.20357.

-
- [141] School of Physics, “Topic 5: Noise in Images,” *The University of Edinburgh*, 2007.
- [142] MathWorks, “Add noise to image - imnoise,” *MathWorks*, 2006. <https://uk.mathworks.com/help/images/ref/imnoise.html> (accessed Jan. 24, 2022).
- [143] F. N. Fritsch and R. E. Carlson, “Monotone Piecewise Cubic Interpolation,” *SIAM Journal on Numerical Analysis*, vol. 17, pp. 238–246, 1980.
- [144] D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*. Upper Saddle River, NJ: Prentice Hall, 1988.
- [145] 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>.
- [146] 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 - Pending Review,” *Journal of Imaging Science and Technology*, vol. 65, no. 6, pp. 60405-1-60405–12, 2022.
- [147] 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.
- [148] L. Torrey and J. Shavlik, “Transfer Learning,” in *Handbook of Research on Machine Learning Applications*, 1st ed., E. Soria, J. Martin, R. Magdalena, M. Martinez, and A. Serrano, Eds. Hershey: IGI Global, 2009, pp. 242–264.
- [149] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-fei, “ImageNet: a Large-Scale Hierarchical Image Database,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2009, doi: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848).
- [150] C. Szegedy *et al.*, “Going deeper with Convolutions,” *Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2015.
- [151] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *International Conference on Learning Representations (ICLR), Computer Vision*, pp. 1–14, 2015.
- [152] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *Computer Vision and Pattern Recognition (CVPR)*, pp. 1–12, 2015.

-
- [153] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places : A 10 million Image Database for Scene Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1452–1464, 2018, doi: 10.1109/TPAMI.2017.2723009.
- [154] J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba, "SUN Database: Large-scale Scene Recognition from Abbey to Zoo," *IEEE Conference on Computer Vision and Pattern Recognition*, 2010. <https://groups.csail.mit.edu/vision/SUN/> (accessed May 18, 2020).
- [155] W. S. Geisler and J. S. Perry, "Statistics for optimal point prediction in natural images," *Journal of Vision*, vol. 11, no. 2011, pp. 1–17, 2016, doi: 10.1167/11.12.14.Introduction.
- [156] University of Texas at Austin, "Natural Scene Statistics in Vision Science Image Database," 2020. <http://natural-scenes.cps.utexas.edu/db.shtml> (accessed May 18, 2020).
- [157] H. Li, C. Yan, and J. Shao, "Measurement of the modulation transfer function of infrared imaging system by modified slant edge method," *Journal of the Optical Society of Korea*, vol. 20, no. 3, pp. 381–388, 2016, doi: 10.3807/JOSK.2016.20.3.381.
- [158] M. Orfanidou, S. Triantaphillidou, and E. Allen, "Predicting image quality using a modular image difference model," *Image Quality and System Performance V*, 2008, doi: 10.1117/12.766407.
- [159] H. Elsaid, G. Thomas, and D. Williams, "Texture features based on the use of the hough transform and income inequality metrics," *Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 1–6, 2017, doi: 10.1109/IPTA.2017.8310094.
- [160] P. Gupta, C. G. Bampis, Y. Jin, and A. C. Bovik, "Natural scene statistics for noise estimation," *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 85–88, 2018, doi: 10.1109/SSIAL.2018.8470313.
- [161] M.-J. Chen and A. C. Bovik, "No-Reference Image Blur Assessment using Multiscale Gradient," in *Quality of Multimedia Experience*, 2009, pp. 70–74, doi: 10.1109/QOMEX.2009.5246973.
- [162] T. O. Aydin, A. Smolic, and M. Gross, "Automated aesthetic analysis of photographic images," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, pp. 31–42, 2015, doi: 10.1109/TVCG.2014.2325047.
- [163] S. Eddins, "Tips for reading a camera raw file into MATLAB," *MathWorks*, 2011. <https://blogs.mathworks.com/steve/2011/03/08/tips-for-reading-a-camera-raw-file-into-matlab/> (accessed Sep. 08, 2021).

