Journal of Imaging Science and Technology® 63(6): 060406-1–060406-11, 2019.

Validation of Modulation Transfer Functions and Noise Power Spectra from Natural Scenes

Edward W. S. Fry
PhD Student, Faculty of Science and Technology, The University of Westminster, UK
E-mail: e.fry@my.westminster.ac.uk

Sophie Triantaphillidou
Reader in Imaging Science, Faculty of Science and Technology, The University of Westminster, UK

Robin B. Jenkin
Principal Image Quality Engineer, NVIDIA Corporation, USA

John R. Jarvis
Visiting Professor of Imaging Science, Faculty of Science and Technology, The University of Westminster, UK

Ralph E. Jacobson
Emeritus Professor of Imaging Science, Faculty of Science and Technology, The University of Westminster, UK

Abstract. *The Modulation Transfer Function (MTF) and the Noise Power Spectrum (NPS) characterize imaging system sharpness/resolution and noise, respectively. Both measures are based on linear system theory. However, they are applied routinely to scene-dependent systems applying non-linear, content-aware image signal processing. For such systems, MTFs/NPSs are derived inaccurately from traditional test charts containing edges, sinusoids, noise or uniform luminance signals, which are unrepresentative of natural scene signals. The dead leaves test chart delivers improved measurements from scene-dependent systems but still has its limitations. In this article, the authors validate novel scene-and-process-dependent MTF (SPD-MTF) and NPS (SPD-NPS) measures that characterize (i) system performance concerning one scene, (ii) average real-world performance concerning many scenes or (iii) the level of system scene dependency. The authors also derive novel SPD-NPS and SPD-MTF measures using the dead leaves chart. They demonstrate that the proposed measures are robust and preferable for scene-dependent systems to current measures.* © *2019 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2019.63.6.060406]

1. INTRODUCTION

This article is concerned with the characterization of the spatial imaging performance of capturing systems, as implemented during their design and manufacture. The *Modulation Transfer Function* (MTF) and the *Noise Power Spectrum* (NPS) are commonly used for such purposes and are principal input parameters to image quality metrics (IQM) from the field [1–4]. The MTF and the NPS measure system signal transfer (relating to sharpness and resolution attributes [2]) and noise, respectively. They are derived traditionally by capturing test charts that provide well-characterized input signals. MTFs and NPSs aim to describe the average real-world performance of the system (i.e. its "general" performance when capturing natural scenes).

The MTF and NPS rely on linear system theory. Strictly speaking, they should only be measured from linear, spatially invariant and homogeneous systems [5]. In theory, they fully specify these systems (Figure 1a) and measurements obtained using any test chart, s_n, are representative of the average real-world system performance, $F(s)$.

However, measurements from systems applying non-linear content-aware image signal processing (ISP), such as noise reduction and sharpening, are dependent on the input signal content. Such "scene-dependent" capture systems are increasingly common, especially in camera phones and autonomous vehicles. Characterizing these systems is far more challenging. Fig. 1(b) defines their average real-world performance, $\overline{F}(s)$, as the mean of all MTFs/NPSs corresponding to the infinite number of potential natural scenes that the system may capture, s_n.

Classical methods for deriving the MTF employ test charts consisting of sinusoids [5–9], edges [5, 9, 10] or random noise signals [11]. These charts are not representative of natural scene signals. As a result, each chart yields significantly different MTFs for non-linear capturing systems [12], which are unrepresentative of average real-world system signal transfer. For the same reason, NPSs derived by traditional methods from uniform luminance patches are unrepresentative of the average real-world noise performance of such systems [13].

IS&T Members.

Received July 21, 2019; accepted for publication Oct. 13, 2019; published online Jan. 7, 2020. Associate Editor: Zeev Zalevsky.
1062-3701/2019/63(6)/060406/11/$25.00

The *dead leaves* test chart was designed with the aim of triggering non-linear content-aware ISPs at comparable levels to natural scenes. It provides a more appropriate signal for measuring non-linear system signal transfer [14–16] and noise [13] and represents a step toward measuring performance using natural scenes. It simulates natural scene textures using a stochastic model [17]. It models the inverse power function of the "average" natural scene, among other natural scene statistics [14].

To date, the NPS has not been measured directly from dead leaves signals. But noise has been measured by comparing MTFs from different dead leaves measurement implementations [13]. Such noise measures have not been used as input parameters to the *direct dead leaves MTF implementation* [15] and should be more appropriate than the currently used uniform patch NPS.

No measurement from any current test chart can represent the MTF/NPS of a non-linear system with respect to a given input scene. Such an MTF/NPS is defined by $F(s_1)$ in Fig. 1(b) if s_1 is the signal of the scene. Branca et al. have made a first attempt to derive MTFs from natural scenes with some success [18]. But no prior art has combined MTFs/NPSs derived from many scenes, $F(s_n)$, to yield an average real-world system performance measure, $\overline{F}(s)$, that accounts for scene dependency.

This article proposes *scene-and-process-dependent MTF (SPD-MTF) and NPS (SPD-NPS)* measurement frameworks that account for the effect of input image content on non-linear content-aware spatial ISPs. There are four SPD-MTF measures and four SPD-NPS measures, which use different input signals. All measures are validated by evaluating measurements from simulated linear and non-linear image capture pipelines. Each measure tests one of the following hypotheses:

1. The NPS can be derived directly from dead leaves signals and is more appropriate than the uniform patch NPS.
2. The accuracy of the direct dead leaves MTF [15] Eq. (3) improves if it accounts for system noise using #1.
3. The MTF/NPS can be derived with respect to a given input scene, accounting for system scene dependency.
4. The average real-world performance of a system can be characterized, accounting for its scene dependency, as the mean of #3 over a large and representative image set.
5. The level of system scene dependency can be measured as the standard deviation of #3 over the same image set.

The following sections of this article define the MTF and NPS, describing current measurement methods and their limitations. The SPD-MTF and SPD-NPS measures are then presented, including their sources of error. The test image dataset and simulated pipelines are described later. We then validate each measure, analyzing sources of error and scene-dependent pipeline behavior. Finally, we draw conclusions on the validity of our hypotheses, the proposed measures and their broader application.

2. BACKGROUND ON MTF AND NPS MEASURES

2.1 *The Noise Power Spectrum*

The NPS characterizes the power of noise introduced by the system, with respect to spatial frequency, u. The one-dimensional (1D) NPS can be obtained from digital systems as the radial average of a two-dimensional (2D) NPS (Eq. (1)) computed using the discrete Fourier transform (DFT). u and v are horizontal and vertical spatial frequencies, respectively. $H(x, y)$ is a noise image of dimensions M by N, expressed by Eq. (2), where $g(x, y)$ is the output image intensity and $\overline{g}(x, y)$ is the expected (or mean signal) value.

$$\mathrm{NPS}(u, v) = \left| \sum_{x=\frac{M}{2}+1}^{M/2} \sum_{y=\frac{N}{2}+1}^{N/2} H(x, y) e^{-2\pi i(ux+vy)} \right|^2 \tag{1}$$

$$\text{where } H(x, y) = g(x, y) - \overline{g}(x, y). \tag{2}$$

When characterizing image capture systems, the NPS is normally derived from captured uniform luminance patches. This renders $\overline{g}(x, y)$ approximately constant at all coordinates if lens shading correction is applied, or patches are captured where lens shading is minimal. However, noise in captured uniform patches is not representative of noise in captured scenes. This comes from the fact that, for all capture systems, photon noise is a function of scene intensity. Moreover, the uniform patch NPS measure fails more noticeably for systems applying non-linear content-aware denoising and sharpening.

Linear denoising algorithms "average-out" noise at the expense of image detail, texture and edge contrast. Non-linear content-aware algorithms adjust their denoising intensity to preserve valuable signal content. Many spatial domain examples apply thresholding in the presence of luminance gradients [19, 20]. Machine learning can also be implemented to denoise certain features more than others [21–25]. Other algorithms operate on a patch-wise level [26], finding "windows" containing similar structure and averaging them to remove noise [27, 28]. Structural signals impede the local removal of noise by all these non-linear algorithms. This renders system noise scene-dependent, as shown in Figure 2 [3, 13]. Uniform patches provide ideal input conditions for these algorithms. When captured, they are generally less noisy than real scenes. Thus, NPSs derived from them often underestimate the average real-world system noise power. Non-linear content-aware sharpening algorithms cause further scene dependency. They amplify the contrast of edges, detail and high frequencies selectively—as described in the next section—increasing noise in these regions more than others.

The SPD-NPS measurement framework proposed in this article aims to account for the above scene-dependent behavior.

2.2 *The Modulation Transfer Function*

The MTF and the related *Spatial Frequency Response* (SFR) are signal transfer measures. They characterize the reproduction of modulation with respect to spatial frequency [5]. The

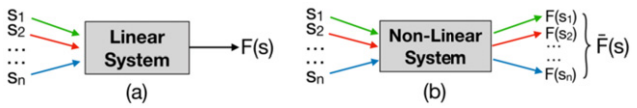

Figure 1. Illustration of theoretical properties of the MTF or NPS ($F(s)$) for signals (s) present in a range of n natural scenes (or test charts): (a) linear system; (b) non-linear system. $F(s)$ is the mean of $F(s_1)$ to $F(s_n)$.

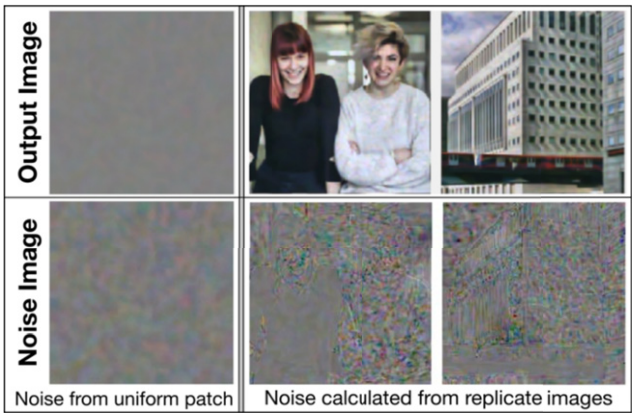

Figure 2. Noise images generated from a non-linear camera simulation pipeline at SNR 10 [3] for the following: uniform patch (left), Branca et al. [18] "People" image (center) and "Architecture" image (right). Noise image contrast was increased to emphasize scene-dependent noise variation.

MTF is obtained from capture systems by three methods, relating to noise [11], edges [9] and sine-wave [5–7, 9] signals. There are various implementations of each method that use particular test charts and processing to deliver the MTF. Each implementation has different sources and levels of measurement error. These include variation error and bias resulting from (i) inaccurate specification of the input signal, (ii) inaccurate measurement of the output signal and (iii) errors resulting from processing when computing the MTF (e.g. DFT computation). Consequently, although a unique MTF exists for linear systems in theory (Fig. 1a), in practice, each implementation produces a different result.

The differences between MTFs derived by these different methods and implementations are far greater, however, for cameras that apply non-linear content-aware denoising and sharpening ISPs. Denoising removes edges, detail and texture. Thus, the mentioned adaptive characteristics of non-linear denoising algorithms also render system signal transfer scene dependent. Non-linear content-aware sharpening introduces further signal transfer scene dependency. For example, adaptive unsharp masks [29–32] reduce contrast amplification in areas of a low signal gradient to avoid boosting noise. Other content-aware sharpening filters operate in various domains [33–36] and can employ guidance images [35] or multi-scale contrast manipulation [36].

Edges, sinusoidal signals and noise have little relation to the average pictorial scene. Their interaction with non-linear content-aware ISPs means that they yield biased MTFs. Namely, the derived MTFs consistently overestimate or underestimate the average real-world performance of non-linear systems, shown as $\bar{F}(s)$ in Fig. 1(b). Edges are generally denoised less and sharpened more than natural scene signals; sinusoidal signals respond less to sharpening [12]. MTFs derived from such signals also fail to describe texture loss [12, 14–16] that is a primary driver of perceived image quality in non-linear capture systems [37]. In order to characterize the real-world performance of non-linear cameras more effectively, MTFs should be measured using test charts that are representative of natural scene signal properties, or even better from scenes themselves, provided this does not significantly increase measurement error.

MTFs are nowadays commonly measured from non-linear systems using the dead leaves chart [14–16] and characterize average real-world system performance and texture loss more effectively. This chart relates more closely to natural scene signals than edges, sinusoidal signals and noise.

The direct dead leaves implementation for measuring the MTF [15] is widely adopted, and core to the IEEE P1858 texture acutance metric and multivariate IQM [2]. It is defined using Eq. (3). It compensates for the system noise power, $\mathrm{NPS}_{\mathrm{Output}}(u)$, using the uniform patch NPS. $\mathrm{PS}_{\mathrm{Input}}(u)$ and $\mathrm{PS}_{\mathrm{Output}}(u)$ are input and output test chart power spectra, respectively and u is the spatial frequency.

$$\mathrm{MTF}(u) = \sqrt{\frac{\mathrm{PS}_{\mathrm{Output}}(u) - \mathrm{NPS}_{\mathrm{Output}}(u)}{\mathrm{PS}_{\mathrm{Input}}(u)}} \tag{3}$$

The recently standardized *intrinsic dead leaves MTF implementation* [16] calculates signal transfer with respect to the cross-spectrum. It is less susceptible to bias from noise since it measures the transfer of both image amplitude and phase content [13]. However, "reversible" image processing, such as sharpening or contrast stretching, is not accounted for [38]. These processes contribute toward both the scene dependency and the perceived image quality of non-linear systems.

Branca et al. derived MTFs with respect to a given natural scene [18] (i.e. $F(s_1)$ in Fig. 1b) by adapting the direct dead leaves implementation (Eq. (3)) so that $\mathrm{PS}_{\mathrm{Input}}(u)$ and $\mathrm{PS}_{\mathrm{Output}}(u)$ are input and output scene power spectra, respectively. This method accounts for signal transfer scene dependency but is often biased since (i) like the direct dead leaves MTF, it compensates for the system's noise, $\mathrm{NPS}_{\mathrm{Output}}(u)$, using the uniform patch NPS and (ii) input images are not windowed or zero-padded, causing periodic replication artifacts that are discussed in the next section. The SPD-MTF measures of this article address these limitations.

3. SCENE-AND-PROCESS-DEPENDENT SYSTEM PERFORMANCE MEASURES

3.1 *Scene-and-Process-Dependent NPSs*

The SPD-NPS measurement framework (Figure 3) derives the NPS using a number of captured images of the same scene or test chart, referred to as *replicates*. The 1D SPD-NPS is defined as the radial average of Eq. (1), where $\bar{g}(x, y)$ is

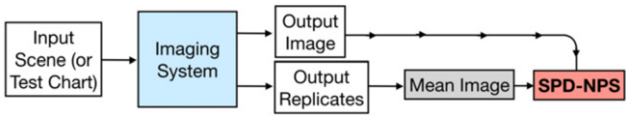

Figure 3. The SPD-NPS measurement framework, adapted from [3].

Figure 4. Artifacts of fixed pattern caused by demosaicing the Branca et al. [18] "People" image (left) and the detail of it (right); image contrast was enhanced.

the mean image of all replicates; other parameters are as previously stated.

The framework accounts for system noise scene dependency but is computationally complex since many replicates must be captured (10 in this article); using fewer replicates underestimates system noise. The framework does not account for demosaicing artifacts of a fixed pattern (Figure 4) or sensor fixed pattern noise. The latter can be measured via other methods [39]. It is also less significant than temporally varying noise in current capture systems under most exposure conditions.

We define the four proposed SPD-NPS measures below.

(i) The *dead leaves SPD-NPS* is derived from the dead leaves chart using the SPD-NPS framework (Fig. 3). It aims to characterize average real-world system noise but requires assumptions of non-linear system behavior.

(ii) The *pictorial image SPD-NPS* is derived from a single pictorial scene using the SPD-NPS framework (Fig. 3). It measures system noise with respect to that scene.

(iii) The *mean pictorial image SPD-NPS* characterizes average real-world system noise, accounting for system scene dependency. It is defined as the mean, $\overline{F}(s)$, of all pictorial image SPD-NPSs, $F(s_n)$, across a set of n scenes, as shown in Fig. 1(b). Averaging NPSs is unorthodox. However, the measure tends toward the average real-world system noise power as n increases, provided that the image set reflects the properties of commonly captured scenes.

(iv) The *pictorial image SPD-NPS standard deviation* measures the level of scene-dependent variation in a system's noise power. It is defined as the standard deviation of all pictorial image SPD-NPSs over a set of n images. Its accuracy increases as n increases.

Measurement bias and variation error in the pictorial image SPD-NPSs are carried into the mean pictorial image SPD-NPS and pictorial image SPD-NPS standard deviation, respectively.

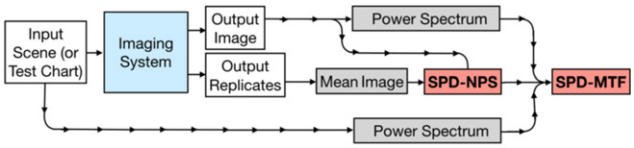

Figure 5. The SPD-MTF measurement framework, adapted from [3].

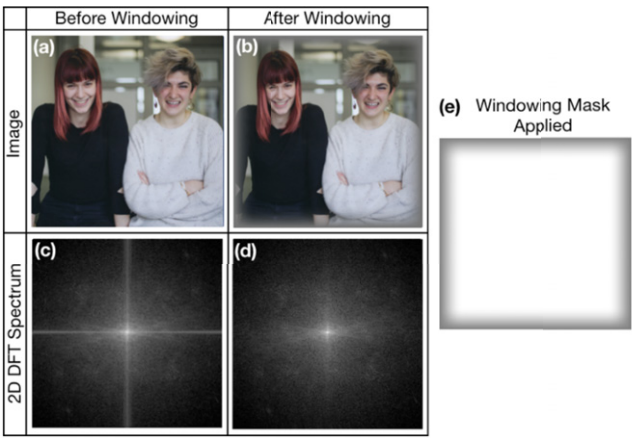

Figure 6. 2D DFT log luminance spectra (c) and (d) for the Branca et al. [18] "People" image (a) and (b), before and after windowing, respectively, with (e).

3.2 *Scene-and-Process-Dependent MTFs*

The SPD-MTF measurement framework (Figure 5) builds upon the previously defined method of Branca et al. [18]. It accounts for scene dependency in terms of system signal transfer and noise. Eq. (3) defines it when $\mathrm{PS}_{\mathrm{Input}}(u)$ and $\mathrm{PS}_{\mathrm{Output}}(u)$ are input and output scene (or dead leaves) power spectra, respectively, and $\mathrm{NPS}_{\mathrm{Output}}(u)$ is the pictorial image (or dead leaves) SPD-NPS.

Input scenes must be windowed to prevent periodic replication artifacts corrupting the 2D luminance spectrum (Figure 6c). These artifacts occur during DFT processing when opposite image edges differ in luminance. They are unaffected by image processing and bias the MTF toward the line $\mathrm{MTF}(u) = 1$. We applied a square-edged mask (Fig. 6e) to taper image edges to a neutral pixel value, starting at 64 pixels from each edge. It uses a cosine function of 1/128 cycles/pixel. It preserves scene signals to mitigate bias from signal-to-noise limitations.

The SPD-MTF inherits signal-to-noise limitations from the direct dead leaves MTF. These are described by Eq. (4). As the input image power, $\mathrm{PS}_{\mathrm{Input}}(u)$, approaches zero at frequency u, $\mathrm{MTF}(u)$ becomes increasingly biased, especially if the measured NPS, $\mathrm{NPS}_M(u)$, underestimates the real system NPS, $\mathrm{NPS}_R(u)$. $\mathrm{PS}_{\mathrm{Output}}(u)$ is the output image power spectrum. Even if the real system NPS is measured with absolute accuracy in a theoretical ideal (i.e. $\mathrm{NPS}_M(u) = \mathrm{NPS}_R(u)$), the numerator and denominator of line 3 of Eq. (4) limit toward zero at equal value and $\mathrm{MTF}(u)$ limits to 1. Thus, test scenes should not have zero (or very low) power at

any frequency. Power spectra should also be computed with satisfactory precision.

$$\begin{aligned}&\text{if } \mathrm{NPS}_M(u) < \mathrm{NPS}_R(u)\\ &\text{then } \lim_{PS_{\mathrm{Input}}(u)\to 0} PS_{\mathrm{Input}}(u) < PS_{\mathrm{Output}}(u) - \mathrm{NPS}_M(u)\\ &\text{and } \lim_{PS_{\mathrm{Input}}(u)\to 0}\left(\frac{PS_{\mathrm{Output}}(u) - \mathrm{NPS}_M(u)}{PS_{\mathrm{Input}}(u)}\right) = \infty\\ &\qquad\qquad = MTF(u).\end{aligned} \tag{4}$$

Signal-to-noise limitations and periodic replication artifacts explain the bias in the scene MTFs of Branca et al. [18]. They also clarify why low-power scenes, and higher frequencies of generally lower power, were most affected. The SPD-MTF framework is less biased since it uses a more appropriate noise measure and windowing. Scenes with significant power across all frequencies, and the dead leaves chart, are expected to yield SPD-MTFs with acceptable bias unless captured under poor exposure conditions. Low-power scenes are still expected to yield biased SPD-MTFs, particularly at higher frequencies, or if captured under poor exposure conditions.

The four proposed SPD-MTF measures are defined below.

(i) The *dead leaves SPD-MTF* characterizes the transfer of dead leaves signals by a system. It aims to describe average real-world signal transfer but requires assumptions of system behavior. It applies the SPD-MTF framework using dead leaves power spectra and the dead leaves SPD-NPS.

(ii) The *pictorial image SPD-MTF* characterizes system signal transfer with respect to a given input scene. It implements the SPD-MTF framework using pictorial scene power spectra and the pictorial image SPD-NPS.

(iii) The *mean pictorial image SPD-MTF* characterizes average real-world system signal transfer, accounting for scene dependency. It is defined as the mean, $\overline{F}(s)$, of all pictorial image SPD-MTFs, $F(s_n)$, across the previously defined set of n scenes (Fig. 1b). Bias is transferred to it from the pictorial image SPD-MTF.

(iv) The *pictorial image SPD-MTF standard deviation* describes the level of system signal transfer scene dependency. It is defined as the standard deviation of all pictorial image SPD-MTFs over the set of n scenes. It is biased by variation error in the pictorial image SPD-MTFs.

4. CAMERA SYSTEM SIMULATION AND TEST IMAGES

Two image capture pipelines were simulated in MATLAB$^{\mathrm{TM}}$. They were tuned to behave similarly to real camera phones at various simulated exposures. Both pipelines modeled physical capture processes and image pre-processing identically, in the following order.

Lens blur was simulated by convolution with a Gaussian approximation for the central lobe of a diffraction-limited lens airy disk, using the f-number and pixel pitch of an *iPhone 6* camera phone [40]. 2D photon noise was simulated using Poisson statistics at maximum linear signal-to-noise ratios (SNR) of 5, 10, 20 and 40 (i.e. at saturation) [41]. It was scaled by factors of 2, 1 and 3.3 in the R, G and B channels, respectively, to account for quantum efficiency variation. Dark current noise was modeled as Gaussian noise with a higher mean and standard deviation at lower SNRs. Black and white levels were adjusted. Pixel information was sampled according to a "grbg" color filter array. Most pipelines simulate the latter process before noise and pre-processing. But the chosen order produced the same output images and facilitated the scaling of photon noise in the R and B channels.

Further, the non-linear pipeline applied the following non-linear content-aware ISPs in order. Demosaicing was by the One Step Alternating Projections (OSAP) [42] algorithm. Denoising was by Block Matching and 3D Filtering (BM3D) [27]. Color channel images were sharpened by the Guided Image Filter (GIF) [35] with no external guidance image specified, and concatenated.

The linear pipeline applied the following equivalent linear ISPs in order. Demosaicing was by the Malvar et al. [43] algorithm. Denoising was by 2D spatial domain Gaussian filtering. Sharpening was by the *imsharpen* MATLAB$^{\mathrm{TM}}$ function.

The image set consisted of 50 high-quality imaged scenes of 512 × 512 pixels, from the LIVE Image Quality Assessment Database [44], Branca et al. [18], Fry et al. [45] and Allen et al. [46] image sets. They reflected the variety of scenes captured by contemporary camera systems.

5. VALIDATION AND DISCUSSION

The SPD-MTF and SPD-NPS measures were validated by analyzing measurements from the linear and non-linear pipelines that are referred to by their Poisson noise SNR. Here, we present results from SNRs 40 and 5, representing very good and very poor signal quality, respectively. SNRs in between showed comparable trends. All NPSs describe luminance noise. Burns's direct dead leaves MTF implementation [47] was adapted for computing all measures. All measurements were smoothed by a moving average filter of seven segments.

5.1 *Scene-and-Process-Dependent NPS*

There is no current way of deriving the ground truth (or "correct") NPS. Nevertheless, Figs. 7 and 8 are used to validate SPD-NPSs derived from the dead leaves chart and pictorial images, respectively. Fig. 7 demonstrates that the dead leaves SPD-NPS (red lines) characterizes the average real-world noise power of the non-linear pipeline more competently than the uniform patch NPS (black line). This validates hypothesis 1. Levels of bias in both measures were comparable since they delivered similar measurements from the linear pipeline. This was as expected, in theory,

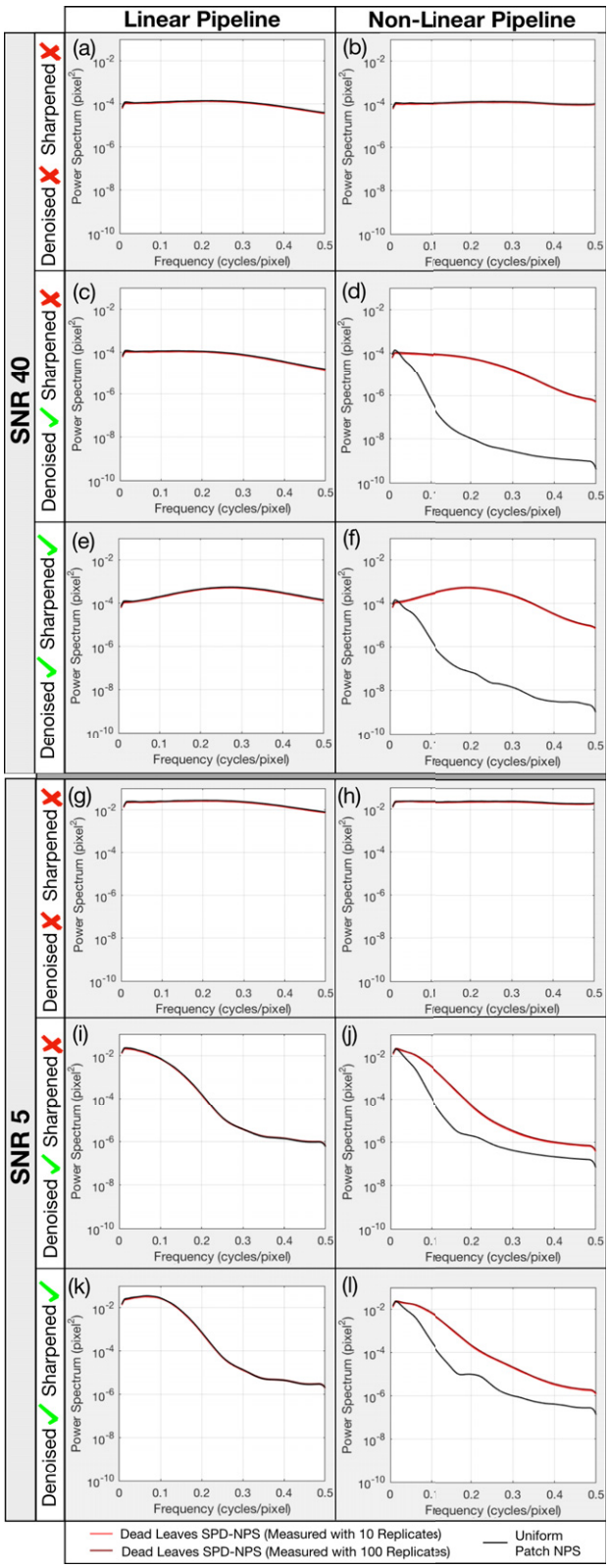

Figure 7. Luminance NPSs derived from the dead leaves chart (red lines) and uniform patches (black lines) at different stages of processing at SNR 40 (a)–(f) and SNR 5 (g)–(l).

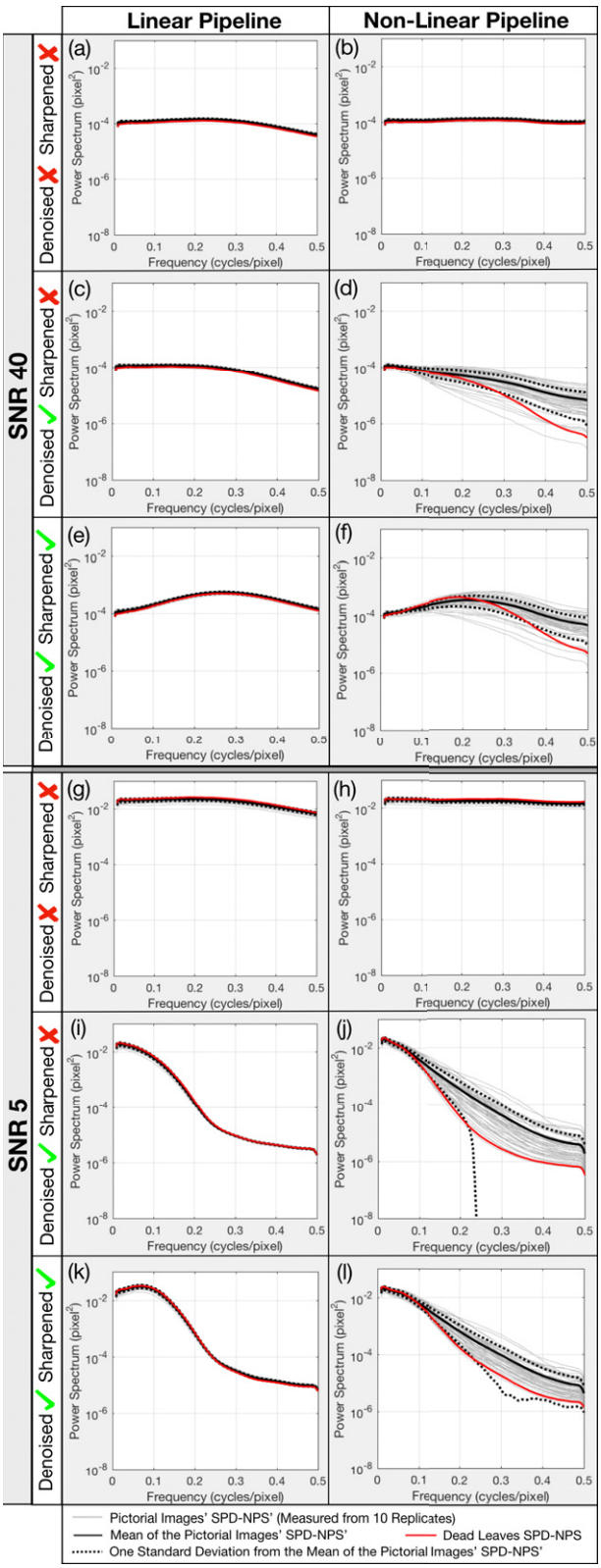

Figure 8. Pictorial image SPD-NPSs (gray lines), mean pictorial image SPD-NPSs (black lines), SPD-NPS standard deviations (black dotted lines) and dead leaves SPD-NPSs (red lines) of luminance noise at different stages of processing at SNRs 40 and 5.

for two NPS measures reliant on linear system theory (Fig. 1a). The uniform patch NPS is derived from a less suitable input signal than the dead leaves SPD-NPS. It should be considered the less representative, thus, the less

"correct" measure. It underestimated the dead leaves SPD-NPS after non-linear denoising (Figs 7(d) and 7(j)). This underestimation was compounded slightly by non-linear sharpening. It was also greater at higher SNRs because mild denoising still cleaned the uniform patch effectively. Dead leaves SPD-NPS measurements derived using 10 and 100 replicates were relatively consistent to one another and difficult to distinguish on logarithmic axes. Thus, SPD-NPSs were derived from pictorial images also using ten replicates, for consistency.

Fig. 8 indicates that measurement bias was similar for the pictorial image SPD-NPS (gray lines), mean pictorial image SPD-NPS (black line) and dead leaves SPD-NPS (red line) since their respective measurements from the linear pipeline were alike. The pictorial image SPD-NPS is validated as the most suitable noise measure for non-linear systems with respect to a given input scene (hypothesis 3). It accounted for scene-dependent system behavior as demonstrated by the variation between measurements from different input scenes after non-linear denoising (Figs. 8(d) and 8(j)).

The mean pictorial image SPD-NPS (Fig. 8, black line) best described the average real-world noise power of the pipelines. It accounted for the general trends in scene-dependent pipeline behavior over the 50 test scenes (hypothesis 4). The dead leaves SPD-NPS (Fig. 8, red line) generally underestimated the SPD-NPS measurements from scenes. This suggests the non-linear content-aware BM3D algorithm (which operates on a patch-wise level) denoised dead leaves signals more effectively than pictorial scenes. This may also apply to other non-linear content-aware denoising filters.

The pictorial image SPD-NPS standard deviation (Fig. 8, black dotted lines) described well the level of noise scene dependency in both pipelines in real-world capture scenarios (hypothesis 5). Non-linear denoising was the primary cause of scene dependency. Non-linear sharpening changed the shape and order of the pictorial image SPD-NPSs (gray lines) in a scene-dependent manner but did not compound their spread. The pictorial image SPD-NPS standard deviation accounts for the latter but not the former. Note that the sudden change in the lower standard deviation boundary in Fig. 8(j) is not a discontinuity. It was caused by the curve crossing the x-axis of a graph with a logarithmic y-axis.

When Fig. 8 was plotted on linear y-axes, we also noted some scene-dependent variation in the pictorial image SPD-NPSs from the linear pipeline. We analyze this variation further in reference [48]. Its standard deviation was around 10% and 15% of the mean pictorial image SPD-NPS at SNRs 40 and 5, respectively. We expect that it is mainly caused by scene-dependent variations in (i) the level of Poisson noise, (ii) the scaling of this noise when simulating color channel quantum efficiency and (iii) the effect of black/white level adjustments. Measurement error scene dependency may also have contributed to these curve variations and could not be distinguished from genuine system scene dependency described above. Regardless of its origin, our simulations suggest this variation should not affect significantly the validity of the pictorial image SPD-NPS, mean pictorial image SPD-NPS or pictorial image SPD-NPS standard deviation.

5.2 *Scene-and-Process-Dependent MTF*

As with the NPS, there is no current method to derive the ground truth (or "correct") system MTF. We validate the SPD-MTFs by comparing them with the direct dead leaves MTF [15]. Figures 9 and 10 support the validation of SPD-MTF measures using the dead leaves chart and pictorial images, respectively. All measurements were heavily biased before denoising at SNR 5 due to signal-to-noise limitations (Eq. (4)). This was because noise power was high at these SNR levels and thus underestimated significantly by all NPS measures. Increasing the number of replicates reduced this bias for each SPD-MTF measure, as demonstrated in Fig. 9 for the dead leaves SPD-MTF. Denoising also mitigated bias in all MTF and SPD-MTF measurements and is generally applied in such situations. Thus, Figs. 9(g) and 9(h) and Figs. 10(g) and 10(h) are hereafter referred to as less relevant conditions.

The dead leaves SPD-MTF (Fig. 9, red lines) uses a more appropriate noise measure for non-linear systems than the direct dead leaves MTF (Fig. 9, black line), as confirmed by observations from Fig. 7. This caused it to characterize non-linear pipeline signal transfer with slightly improved accuracy, under the most relevant conditions (hypothesis 2). Both the above measures have similar levels of bias since measurements from the linear pipeline were alike under such conditions. Reducing the number of replicates from 100 to 10 increased the underestimation of noise by the dead leaves SPD-NPS. This, in turn, introduced minor positive bias to dead leaves SPD-MTFs at higher frequencies due to signal-to-noise limitations (Eq. (4)).

The other SPD-MTF measures were derived from pictorial images, also using ten replicates, for fair comparison (Fig. 10). The pictorial image SPD-MTF (gray lines) is the most valid measure, in theory, for non-linear system signal transfer concerning a given scene; such measurements are particularly relevant to image quality modeling applications. It accounted most comprehensively for system scene dependency (hypothesis 3), showing significantly higher variation after application of non-linear ISPs than linear ISPs.

We note, however, that scene-dependent variation was also present in measurements from the linear pipeline under some conditions. We expect this variation to be mainly caused by bias from signal-to-noise limitations (Eq. (4)). This bias was mitigated by denoising but affected the "correctness" of the measure and should be investigated further. It was also scene-dependent, due to variations in scene signal power, and in the underestimation of noise power by the pictorial image SPD-NPS measure. Higher frequencies of low-power images, at lower SNRs, were most affected. It is not presently possible to distinguish between the effects of this bias and measurement variation resulting from genuine system scene dependency.

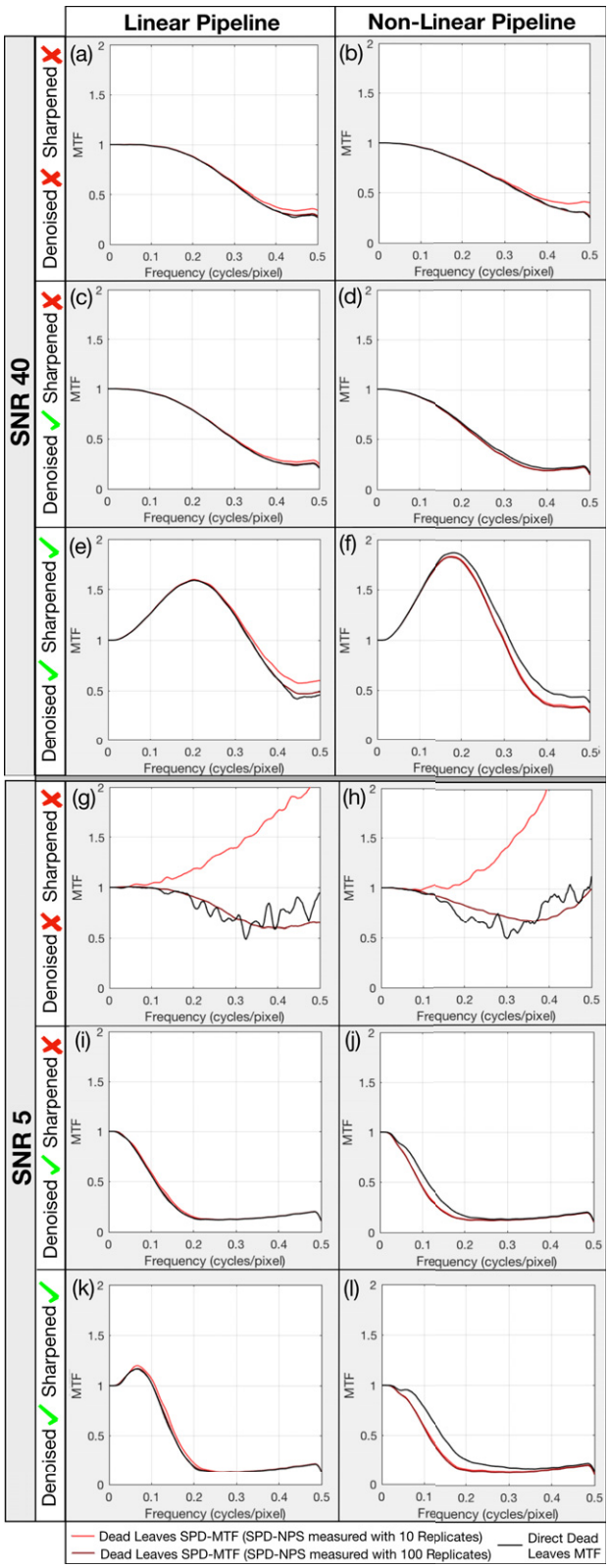

Figure 9. Dead leaves SPD-MTFs (red lines) and direct dead leaves MTFs (black lines) at different stages of processing, at SNR 40 (a)–(f) and SNR 5 (g)–(l). Windowing was not applied for these measurements.

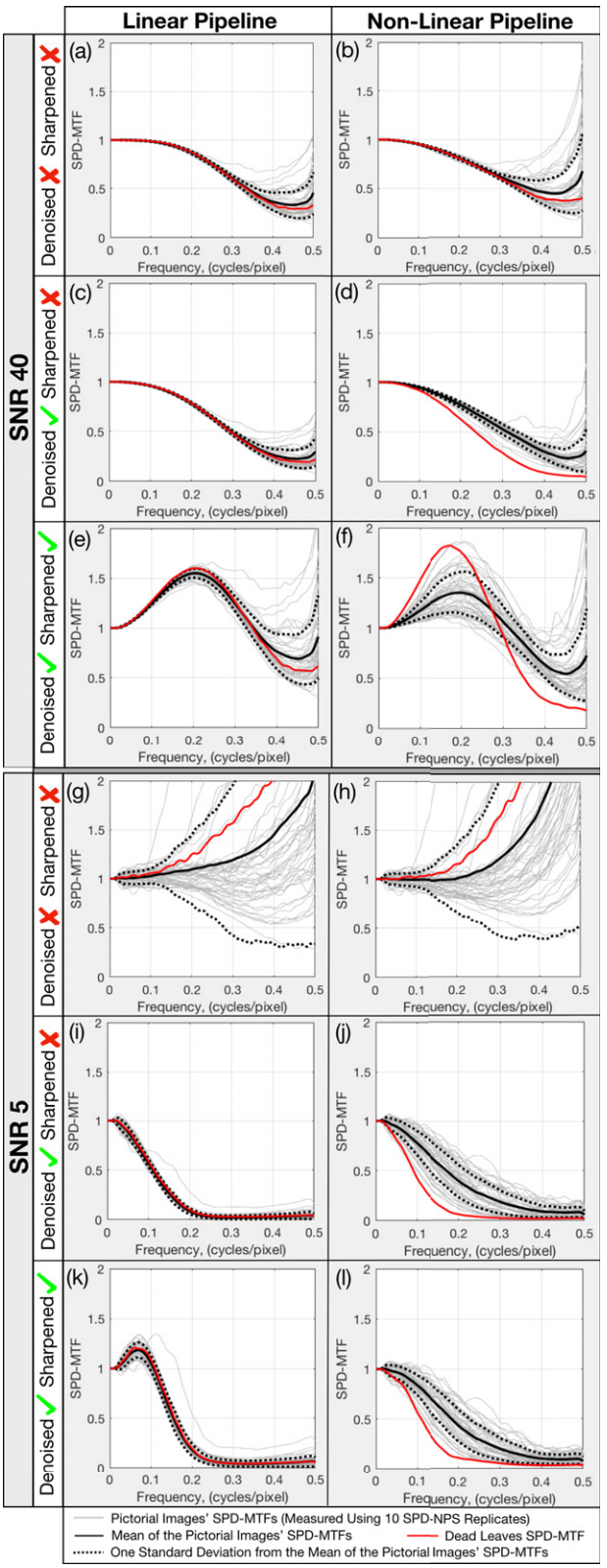

Figure 10. Pictorial image SPD-MTFs (gray lines), mean pictorial image SPD-MTFs (black lines), SPD-MTF standard deviations (black dotted lines) and dead leaves SPD-MTFs (red lines) at different stages of processing at SNRs 40 and 5. All scenes and test charts were windowed for these measurements.

The pictorial image SPD-MTF standard deviation (Fig. 10, black dotted lines) described the overall level of system scene dependency effectively (hypothesis 5). Measurements were greater for the linear pipeline than as

expected in theory. They were also higher proportionally than for the equivalent noise scene dependency measure (Fig. 8, black dotted lines). This was caused by the mentioned scene-dependent variation in pictorial image SPD-MTF measurements. Nevertheless, we expect the pictorial image SPD-MTF standard deviation to be a valuable measure, especially if bias in the pictorial image SPD-MTF can be reduced further. Like the equivalent SPD-NPS measure, it accounts for the spread of the pictorial image SPD-MTF curves but not changes in their shape or order.

The mean pictorial image SPD-MTF (black line) is validated in Fig. 10 as a more relevant measure for average real-world signal transfer in non-linear systems than the dead leaves SPD-MTF (red line). Bias in both measures was similar under the most relevant conditions. This was because calculating the mean of the pictorial images' SPD-MTFs averaged out their scene-dependent variation and bias. This produced curves of similar shape to the dead leaves SPD-MTF for the linear pipeline. This suggests that—provided it is measured from a representative scene set—the mean pictorial image SPD-MTF is often less biased than the individual pictorial image SPD-MTF measurements it is derived from.

After applying non-linear ISPs, however, the dead leaves SPD-MTF was dissimilar to the pictorial image SPD-MTFs and often underestimated the mean pictorial image SPD-MTF. Both of the other measures account for scene dependency more comprehensively than the first. We infer that the non-linear content-aware denoising and sharpening algorithms processed dead leaves signals in a different manner to the average pictorial scene. For a fair comparison, the dead leaves chart was windowed as per all scenes to mitigate bias from periodic replication artifacts.

6. CONCLUSIONS

In this article, we reviewed the limitations of current MTF and NPS measures when characterizing capture systems that apply non-linear, content-aware spatial ISPs. We then proposed several scene-and-process-dependent MTF and NPS measures that account for the scene dependency of such systems. These measures were validated using simulated camera phone capture pipelines. The fact that they revealed significant scene dependency in the non-linear pipeline indicates their promise when characterizing real non-linear cameras and other non-linear imaging systems.

The *pictorial image SPD-MTF* and *SPD-NPS* measured pipeline performance suitably with respect to a given input scene, accounting for scene-dependent behavior. They are the only current measures capable of such characterization. In a separate publication [49], we demonstrate their suitability as input parameters for IQMs when modeling the perceived quality of a given captured scene.

The *mean pictorial image SPD-MTF* and *SPD-NPS* characterized average real-world pipeline performance, accounting for general trends in scene dependency. We propose them as performance optimization parameters and IQM input parameters for applications concerning average real-world system image quality. Current equivalent measures do not account for system scene dependency.

The *pictorial image SPD-MTF* and *SPD-NPS standard deviation* described the degree of scene-dependent variation in each pipeline's performance. They are the only current measures for system scene dependency but do not account for all aspects of it. Combining these measurements with the mean pictorial image SPD-MTF and SPD-NPS characterizes both the average real-world system performance and its level of scene dependency.

The *dead leaves SPD-MTF* and *SPD-NPS* measured non-linear pipeline performance more accurately than the current direct dead leaves MTF and uniform patch NPS, respectively. We propose that they should be used to estimate average real-world system performance. However, results from both these measures were often outliers compared to measurements derived from scenes. This is because non-linear content-aware ISPs were triggered at different levels by the dead leaves chart, compared to natural scene signals. The dead leaves chart simulates an average scene signal with a typical power spectrum. However, with respect to these algorithms, it is only a mathematically generated image with limited relation to the complex spatial signals in pictorial scenes. We propose that a more suitable test chart should either simulate natural scene signal structure more comprehensively or be composed of such signals.

The SPD-NPS measures account for the effect of relevant input signals on temporally varying noise in non-linear systems. They showed little measurement error. Thus, they are more appropriate than current equivalent measures for such systems (if equivalent measures exist). They do not account for fixed patterns of noise/artifacts. Their requirement for many replicates causes all SPD-MTFs and SPD-NPSs to be more computationally complex than current measures.

The SPD-MTF measures are more suitable for non-linear capture systems than current equivalent measures. However, the pictorial image SPD-MTF was prone to measurement bias under certain conditions due to signal-to-noise limitations inherited from the direct dead leaves MTF implementation. This bias was mitigated by denoising. It was also reduced by computing the measure either with more replicates or with higher power scenes. Further investigations are recommended to reduce it further.

This bias was scene dependent and affected the accuracy of the pictorial image SPD-MTF standard deviation. It cannot currently be distinguished from genuine measurement scene dependency that results from interactions between image signals and non-linear ISPs. It averaged itself out over the 50 test scenes, however, to similar levels to the bias in the dead leaves SPD-MTF. Thus, we expect the mean pictorial image SPD-MTF to be less affected than other SPD-MTFs derived from pictorial scenes.

In-depth analysis of this bias and other forms of SPD-MTF and SPD-NPS measurement error were beyond the scope of this article. We suggest that further analysis of such an error is needed to evaluate the "correctness" of each

SPD-MTF and SPD-NPS measure. For the SPD-MTFs, this may involve adapting error propagation methods for existing dead leaves MTF measurement implementations [50].

All the proposed SPD-MTF and SPD-NPS measures would benefit from validation with real capturing systems. In a parallel article [49], we have used them as input parameters to IQMs with the aim of modeling the quality of individual scenes. The modified IQMs were more successful (i.e. they correlated more accurately with perceived image quality). We also developed novel *log Noise Equivalent Quanta (log NEQ)* and *Visual log NEQ* metrics from these measures, which performed very competitively. In a separate publication [48], we have used SPD-NPS measures to examine camera pipeline noise scene dependency in further detail. We also have developed the former into objective single-figure metrics for system noise performance and scene dependency.

REFERENCES

[1] E. M. Crane, "Acutance and granulance," Proc. SPIE Image Qual. **310**, 125 (1981).

[2] IEEE Std 1858^{TM}-2016: IEEE Standard for Camera Phone Image Quality. IEEE Standards Association Board of Governors (IEEE, Piscataway, NJ, 2016).

[3] E. W. S. Fry, S. Triantaphillidou, R. E. Jacobson, J. R. Jarvis, and R. B. Jenkin, "Bridging the gap between imaging performance and image quality measures," *IS&T Electronic Imaging: Image Quality and System Performance XV Proceedings* (IS&T, Springfield, VA, 2018).

[4] P. G. J. Barten, "Evaluation of the effect of noise on subjective image quality," Proc. SPIE **1453**, 2 (1991).

[5] J. C. Dainty and R. Shaw, *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes* (Academic Press Ltd., London, 1974).

[6] J. W. Coltman, "The specification of imaging properties by response to a sine wave input," J. Opt. Soc. Am. A **44**, 468 (1954).

[7] R. L. Lamberts, "Sine-wave response techniques in photographic printing," J. Opt. Soc. Am. A **51**, 982 (1961).

[8] K. R. Castleman, *Digital Image Processing* (Prentice Hall International, 1996).

[9] ISO 12233-2017: Photography — Electronic still picture imaging — resolution and spatial frequency responses (ISO Geneva) www.iso.org .

[10] R. A. Jones, "An automated technique for deriving MTF's from edge traces," Photogr. Sci. Eng. **11**, 102–106 (1967).

[11] P. D. Burns and J. Koplowitz, "Verification of a method to estimate the wiener kernel transforms of a nonlinear system," *Proc. 9th Pittsburgh Conf. on Modelling and Simulation* (Instrument Society of America (ISA), 1978), pp. 1189–1194.

[12] U. Artmann, "Image quality assessment using the dead leaves target: experience with the latest approach and further investigations," Proc. SPIE **9404**, 1–15 (2015).

[13] U. Artmann, "Measurement of noise using the dead leaves pattern," *IS&T Electronic Imaging: Image Quality System Performance XV Proceedings* (IS&T, Springfield, VA, 2018), pp. 341-1–341-6.

[14] F. Cao, F. Guichard, and H. Hornung, "Dead leaves model for measuring texture quality on a digital camera," Proc. SPIE **7537**, 75370E (2010).

[15] J. McElvain, S. P. Campbell, J. Miller, and E. W. Jin, "Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems," Proc. SPIE Digital Photography **Vol. 7537**, 75370D-1 (2010).

[16] ISO 19567-2: 2019: Photography — Digital cameras, Part 2: Texture analysis using stochastic pattern (ISO Geneva), www.iso.org.

[17] G. Matherton, *Random Sets and Integral Geometry* (John Wiley & Sons, New York, 1975).

[18] R. Branca, S. Triantaphillidou, and P. D. Burns, "Texture MTF from images of natural scenes," *IS&T Electronic Imaging: Image Qual. Syst. Perform. XIV Proceedings* (IS&T, Springfield, VA, 2017), pp. 113–120.

[19] J. B. Phillips, S. M. Coppola, E. W. Jin, Y. Chen, J. H. Clark, and T. A. Mauer, "Correlating objective and subjective evaluation of texture appearance with applications to camera phone imaging," *Proc. SPIE-IS&T Electron. Imaging: Image Qual. Syst. Perform. VI* (IS&T, Springfield, VA, 2009), Vol. 7242, p. 724207-1.

[20] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Sixth International Conference on Computer Vision* (IEEE, Piscataway, NJ, 1998), pp. 839–846.

[21] S. Roth and M. J. Black, "Fields of Experts: a framework for learning image priors," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2005).

[22] M. Elad and M. Aharon, "Image Denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process. **15**, 754–758 (2006).

[23] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," IEEE Trans. Image Process. **22**, 700–711 (2013).

[24] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising with multi-layer perceptrons, part 1: comparison with existing algorithms and with bounds," J. Mach. Learn. Res. **13**, 1–38 (2012).

[25] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," *2011 Int'l. Conf. Computer Vision* (IEEE, Piscataway, NJ, 2011), pp. 479–486.

[26] R. C. Sumner, R. Burada, and N. Kram, "The effects of misregistration on the dead leaves cross-correlation texture blur analysis," *IS&T Electronic Imaging: Image Quality and System Performance XIV Proceedings* (IS&T, Springfield, VA, 2017), pp. 121–129.

[27] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," Proc. SPIE **6064**, 1–12 (2006).

[28] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," Multiscale Model. Simul. **4**, 490–530 (2005).

[29] G. Ramponi and A. Polesel, "Rational unsharp masking technique," J. Electron Imaging **7**, 333 (1998).

[30] S. H. Kim and J. P. Allebach, "Optimal unsharp mask for image sharpening and noise removal," J. Electron Imaging **14**, 2 (2005).

[31] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," IEEE Trans. Image Process. **9**, 505–510 (2000).

[32] G. Deng, "A generalized unsharp masking algorithm," IEEE Trans. Image Process. **20**, 1249–1261 (2011).

[33] X. Zhu and P. Milanfar, "Restoration for weakly blurred and strongly noisy images," *2011 IEEE Workshop on Applications of Computer Vision* (IEEE, Piscataway, NJ, 2011), pp. 103–109.

[34] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," ACM Trans. Graph. **30**, 69 (2011).

[35] K. He, J. Sun, and X. Tang, "Guided image filtering," IEEE Trans. Pattern Anal. Mach. Intell. **35**, 1397–1409 (2010).

[36] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *Transactions on Graphics - Proc. ACM SIGGRAPH* (ACM, New York, NY, 2008), p. 67.

[37] E. W. Jin, J. B. Phillips, S. Farnand, M. Belska, V. Tran, E. Chang, Y. Wang, and B. Tseng, "Towards the Development of the IEEE P1858 CPIQ Standard – a validation study," *IS&T Electron. Imaging: Image Quality and System Performance Proceedings* (IS&T, Springfield, VA, 2017), pp. 88–94.

[38] P. D. Burns and J. Martinez Bauza, "Intrinsic camera resolution measurement," Proc. SPIE **9396**, 939609 (2015).

[39] ISO 15739 :2017: *Photography — Electronic still-picture imaging — Noise measurements* (ISO Geneva), www.iso.org.

[40] Apple Inc., "iPhone 6 - Technical Specifications,". Available https://support.apple.com/kb/sp705?locale=en_GB, accessed December 2016.

[41] MathWorks, "Imnoise: Add noise to image,". Available http://uk.mathworks.com/help/images/ref/imnoise.html, accessed April 2018.

[42] Y. M. Lu, M. Karzand, and M. Vetterli, "Demosaicking by alternating projections: Theory and fast one-step implementation," IEEE Trans. Image Process. **19**, 2085–2098 (2010).

[43] H. S. Malvar, L. He, and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images," *2004 IEEE Int'l. Conf. Acoustics Speech Signal Processing* (IEEE, Piscataway, NJ, 2004), pp. 5–8.

[44] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2," 2008. [Online]. Available: http://live.ece.utexas.edu/research/quality, accessed June 2017.

[45] E. W. S. Fry, S. Triantaphillidou, J. Jarvis, and G. Gupta, "Image quality optimization, via application of contextual contrast sensitivity and discrimination functions," Proc. SPIE **9396** (2015).

[46] E. Allen, S. Triantaphillidou, and R. E. Jacobson, "Perceptibility and acceptability of JPEG 2000 compressed images of various scene types," Proc. SPIE **9016** (2014).

[47] P. D. Burns, "Refined measurement of digital image texture loss," Proc. SPIE **8653**, 86530H (2013).

[48] E. W. S. Fry, S. Triantaphillidou, R. B. Jenkin, R. E. Jacobson, and J. R. Jarvis, "Noise power spectrum scene-dependency in simulated image capture systems," *IS&T Electronic Imaging 2020: Image Qual. Syst. Perform. XVII Proceedings* (IS&T, Springfield, VA, 2020), (submitted, under review).

[49] E. W. S. Fry, S. Triantaphillidou, R. B. Jenkin, R. E. Jacobson, and J. R. Jarvis, "Scene-and-process-dependent spatial image quality metrics," J. Imaging Sci. Technol. **63** (2020) (in print).

[50] P. D. Burns, "Estimation error in image quality measurements," Proc. SPIE **7867**, 78670H (2011).