Non-uniformity Correction of IR Images using Natural Scene Statistics

Todd Goodall, Alan C. Bovik, and Haris Vikalo
University of Texas at Austin

Nicholas G. Paulter, Jr.
National Institute of Standards and Technology

***Abstract*—Infrared images are commonly afflicted by distortions such as non-uniformity. Non-uniformity is characterized by horizontal and vertical fixed pattern noise. Accurately estimating the amount of non-uniformity present in an image and removing that amount of non-uniformity noise are open problems. Several estimators of non-uniformity exist, but their ability to estimate degrades with the presence of other sources of noise. Specifically, most of these metrics lack the robustness demanded by a more complete non-uniformity model.**

Previous non-uniformity correction algorithms are compared and found to underperform relative to a more complete model of non-uniformity that we have developed. Using this model, we have created a new denoising algorithm, which we call the Gaussian scale mixture perceptual pattern denoiser. The new model and algorithm can fully characterize non-uniformity using covariance matrices.

***Index Terms*—Non-uniformity; natural scene statistics; fixed-pattern noise; Gaussian scale mixture; non-uniformity correction**

I. Introduction

Long wavelength infrared (LWIR) images are often used for target detection and temperature measurement. Capturing wavelengths between 8 and 14 micrometers, these images have the ability to "see" through smoke, fog, and dust and are especially useful in low visible-light conditions. For example, firefighters image the heat in a room to determine critical burn points and environmental hazards. The ability of a firefighter to detect such points and hazards is ultimately connected to saving lives [1]. Other applications of infrared images include surveillance, missile guidance, and night vision.

Given the many uses of LWIR images, they have been well studied. Mooney characterized sources of spatial noise [2] and the effect of noise on minimum resolvable temperature differences (MTD) as a function of frequency [3]. Lopez-Alonso further characterized spatial noise in IR images by using Principle Components Analysis (PCA) to separate spatial and temporal noise from a sequence of frames [4]. These studies led Pezoa and Medina to model the non-uniformity (NU) noise using spectral domain characteristics [5] that are distinctly different from independent (white) spatial noise. Using this NU model, Pérez *et al.* measured and compared the efficacy of several multi-frame non-uniformity correction (NUC) algorithms [6] and developed methods for extracting the structure of the underlying fixed-noise pattern [7].

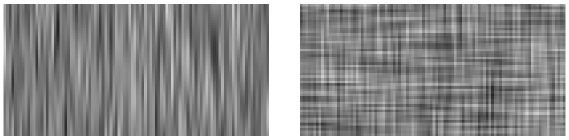
This study and characterization of noise behavior in infrared images motivates further study regarding noise removal. To date, several effective solutions exist for reducing additive white noise in visible light images. More recently, the Gaussian scale mixture model (GSM) has been successfully used in the development of algorithms that can remove white noise in images by utilizing the fact that wavelet subbands can be "Gaussianized" by multiplication by a random variable [8]. In LWIR images, additive noise is often a mixture of white noise and non-uniformity noise. For additive white noise, Kafieh and Rabbani apply locally adaptive wavelet shrinkage successfully to medical infrared images [9]. A search of the available literature yields little information regarding the utility of wavelet denoising techniques for non-uniformity within infrared images.

Top-performing LWIR non-uniformity denoising algorithms have generally been designed using heuristics. In addition, denoising algorithms are usually based on an oversimplified assumption of the non-uniformity. Total variation non-uniformity correction (TVNUC) [10] performs iterative and weighted smoothing to reduce the "roughness" as measured by the roughness index. Another top-performing algorithm, the midway infrared equalization (MIRE) [11], corrects column-wise histograms of an image by locally averaging empirical column-wise cumulative distribution functions (CDFs) for removing spatial flicker. As will be explained later, a more accurate model of non-uniformity noise can be described as a well-defined spatially-varying striping pattern with certain spectral domain properties.

Moreover, these algorithms do not take into account the powerful priors presented by a deeper understanding of LWIR natural scene statistics (NSS). Such a study found that NSS can be used to evaluate LWIR image quality [12]. Additionally, this study provided evidence that NSS features can produce a highly effective estimate for detecting the presence and magnitude of non-uniformity noise in infrared images. These statistics allow a NU correction algorithm to gauge the quality of an image before beginning the correction process to achieve an near-optimal result, as will be presented in the proceeding sections.

II. Images

A corpus of 468 high-quality LWIR images was built up from the NIST [13] and MORRIS [14] databases. The NIST database contains images taken from an indoor setting



(a) I_{NU} (columns) (b) I_{NU} (columns and rows)

Fig. 1. Model of NU in infrared images

containing many hotspots, and each 16 bit image has resolution 640x480. The MORRIS database contains mostly urban images of people, cars, and buildings, and each 8 bit image has resolution 384x288.

Each image is converted to floating point and normalized so that the range of luminances is between 0 and 1. This allows for simplifying calculations regarding noise magnitude since the bit depths between NIST and MORRIS differ.

III. Non-uniformity Models

Non-uniformity is produced by combinations of vertical or horizontal striping noise that arises from manufacturing inconsistencies in the read-out architecture found in infrared focal-plane sensor arrays. These patterns may also result from dark current and segmented sensor capture areas [5] [15] [4].

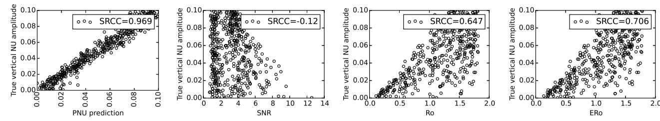
Often, this striping artifact is assumed to only produce uniform stripes aligned along one dimension of an image. Unfortunately, this assumption is not entirely accurate. Pezoa and Medina proposed another model of non-uniformity after studying its spectral properties [5]. They found that the horizontal and vertical lines that appear in infrared images are not simply independent gains. They define an additive noisy non-uniformity image I_{NU} in the spectral domain using

$$|\tilde{I}_{NU}(u,v)| = B_u \exp\left(\frac{-\left(u-u_0\right)^2}{2\sigma_u^2}\right) + B_v \exp\left(\frac{-\left(v-v_0\right)^2}{2\sigma_v^2}\right) \quad (1)$$

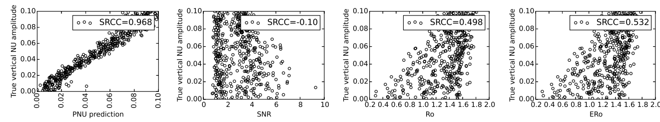
$$\angle \tilde{I}_{NU}(u,v) \sim \mathrm{U}[-\pi,\pi] \quad (2)$$

where $\tilde{I}_{NU}$ is the Fourier Transform representation of the noise image I_{NU}, $B_u = B_v = 5.2$, $\sigma_u = \sigma_v = 2.5$, and where $\mathrm{U}[a,b]$ denotes the uniform distribution on $[a,b]$. The severity of NU can be controlled by scaling the dynamic range of I_{NU} using a standard deviation parameter σ. I_{NU} is scaled to match the dynamic range of a same-sized matrix sampled from $\mathcal{N}(0,\sigma^2)$ and added to pristine image to produce the distorted image. This model will be referred to as the spectral non-uniformity model. Column-based spectral NU is exemplified in Fig. 1(a) (by letting $B_u = 0$), and the full spectral NU model is depicted in Fig. 1(b).

In NUC algorithms, producing a reference-free estimate of the NU in an image is essential to evaluating correction performance on-the-fly [6]. Commonly used methods for estimating NU magnitude include the Roughness index (Ro),



(a) PNU (b) SNR (c) Ro (d) ERo



(e) PNU (f) SNR (g) Ro (h) ERo

Fig. 2. Scatter-plots comparing ground truth vertical NU standard deviation (σ_V) to model predictions in Figs. 2a,2b,2c, and 2d. Predictions made on images also distorted with additive white noise in Figs. 2e, 2f, 2g, and 2h.

Effective Roughness index (ERo), and SNR. Since LWIR images commonly contain both non-uniformity noise and additive white noise, measurements of non-uniformity noise should be largely independent of noise, and vice versa.

A common method for estimating NU is the spatial signal-to-noise ratio (SNR) of the image, I, using μ/σ where σ and μ are the standard deviation and mean pixel intensities within a localized area. Another common method for estimating NU is the Roughness [16] index:

$$\mathrm{Ro}(I) = \frac{\|h_1 * I\|_1 + \|h_2 * I\|_1}{\|I\|_1} \quad (3)$$

where h_1 is the 1-D differencing filter with impulse response $[1,-1]$, $h_2 = {h_1}^T$, $\|\cdot\|_1$ is the L_1 norm, and $*$ indicates convolution. Based on Ro, the Effective Roughness [17] index is then defined as

$$\mathrm{ERo}(I) = \mathrm{Ro}(g * I) \quad (4)$$

where g is a high-pass filter, with the additional modification that in (4), the L_2 norm is used in (3) instead of the L_1 norm. Here $g * I = I - \mu$.

Each of these roughness indices measures the general roughness of a image, but they do not account for the natural scene statistics properties found to exist in both visible light and LWIR images. A novel method for measuring NU, the perceptual NU index (PNU), was developed in [12], where 102 features are extracted from an image and mapped to a ground truth σ_{NU} value as described in the spectral noise model. This method uses machine-learning to avoid some weaknesses found in other approaches. Observed weaknesses with other NU estimators include dependence on the size of the input image, sensitivity to high frequency content in an image, and dependence on other types of noise in an image.

To test the predictive capabilities of each of the NU estimators, the image corpus was distorted with simulated spectral vertical NU and horizontal NU of parameters $\sigma_V, \sigma_H \in [0, 0.1]$. Given this continuous range, PNU is trained on ground truth vertical NU magnitude using a leave-one-out policy (i.e. one content type was left out and the remaining content used for testing). Also, Ro and ERo indices were appropriately modified to focus on vertical distortion. The estimates for σ_V are plotted in Fig. 2 for SNR, Ro, ERo, and PNU indices. From Fig. 2(a), PNU is observed to be very highly correlated (Spearman rank Correlation Coefficient (SRCC) of 0.97) with σ_V. The SNR estimate in Fig. 2(b) appears uncorrelated with σ_V. Ro and ERo in Figs. 2(c) and 2(d) demonstrate a weak correlation. Clearly, the best estimator is PNU since it maintains a clearer one-to-one relationship between prediction and ground truth and is robust against sources of horizontally striping non-uniformity.

To test this idea in the presence of additive white noise (AWN), the same test as above was duplicated with the exception that each image was also randomly distorted with AWN distributed as $\mathcal{N}(0, \sigma_{WN}^2)$ where standard deviation, $\sigma_{AWN} \in [0, 0.1]$. The same NU estimates are computed as before, and corresponding scatter-plots are depicted in the bottom row in Fig. 2 for the SNR, Ro, ERo, and PNU indices. Ro and ERo indices appear to further decorrelate from σ_V, losing a significant portion of their predictive capabilities with the additional noise. However, PNU is robust and measures almost as well as the case in Fig. 2(a) without AWN. Again, the best overall estimator is PNU as it is largely unaffected by the AWN.

IV. GSMPP Model

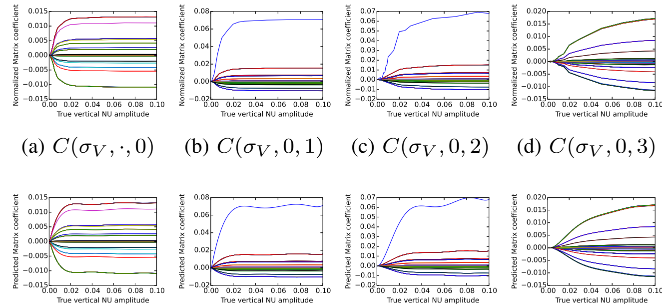
Newly proposed here is a GSM-based perceptual pattern (GSMPP) denoiser. Inspired by the perceptual aspect and success of the gaussian scale mixture model, GSMPP provides a theoretical model for removing the pattern noise that has been noted to afflict infrared images.

Previously, Gaussian scale mixture (GSM) [8] denoising has been applied to visible light images perturbed by white noise. The GSM denoiser allows wavelet coefficients to be "Gaussianized" then well characterized by their second moments computed with respect to neighboring subband coefficients. Possibly due to the perceived difficulty in extracting noise covariance matrices without knowing the image content, this method has not previously been extended further for noise that introduces dependencies and patterns.

The GSM model is applied on the result of a steerable pyramid decomposition [18], which separates an image into 6 orientations, 3 scales, a high-pass residual band, and a low-pass residual band. From this transform, a random vector $\bar{g}(\mathbf{x}, \theta, b)$ is gathered with respect to spatial location $\mathbf{x} = (x, y)$, orientation θ, and pyramid level b. Let $b = 0$ refer to the high-pass residual, $b = \{1, 2, 3\}$ refer to the scale-orientated subbands, and $b = 4$ refer to the low-pass residual. This vector, for $b \in \{0, 1, 2\}$ is constructed by collecting 9 coefficients plus 1 at each of the 5 neighboring orientated subbands, and 1 from the parent subband to produce a total length (M) of 15. Vectors from $b = 3$ are constructed the same except without the parent subband coefficient to produce a total length (M) of 14.

These extracted vectors, $\bar{g}$ are assumed to follow the Gaussian scale mixture model

$$\bar{g}(\mathbf{x}, \theta, b) = z(\mathbf{x}, \theta, b)\bar{\gamma}(\mathbf{x}, \theta, b) \tag{5}$$



(a) $C(\sigma_V, \cdot, 0)$ (b) $C(\sigma_V, 0, 1)$ (c) $C(\sigma_V, 0, 2)$ (d) $C(\sigma_V, 0, 3)$

(e) poly-fit of 3a (f) poly-fit of 3b (g) poly-fit of 3c (h) poly-fit of 3d

Fig. 3. Plots 3a through 3d represent sampled covariance matrix coefficients for different σ_V. Plots 3e through 3h represent the 6th order poly-fits used to approximate the true noise covariance matrix coefficients as a function of σ_V. All plots contain $15 * 15 = 225$ coefficients except for plots 3d and 3h which represent $14 * 14 = 196$ coefficients. Most lines overlap indicating matrix coefficients that mutually scale at the same rate with σ_V.

where z is a positive random scalar estimated, using maximimum likelihood, to be

$$\hat{z}_{ML} = \sqrt{\frac{\bar{g}^T C_\gamma^{-1} \bar{g}}{M}} \tag{6}$$

where $\bar{\gamma} \sim \mathcal{N}(0, C_\gamma)$. Then if an image is contaminated by an additive noise N, we have

$$\bar{g}_N(\mathbf{x}, \theta, b) = \bar{g}(\mathbf{x}, \theta, b) + N(\mathbf{x}, \theta, b). \tag{7}$$

The best mean-squared-error (MSE) estimate of the true value of $\bar{g}(x, y, \theta, b)$ given $\bar{g}_N(x, y, \theta, b)$ and $z(x, y, \theta, b)$ is

$$\begin{aligned} \hat{g}(\mathbf{x}, \theta, b) &= E\left[\bar{g}(\mathbf{x}, \theta, b) | \bar{g}_N(\mathbf{x}, \theta, b), z(\mathbf{x}, \theta, b)\right] \\ &= \hat{z} C_{\bar{\gamma}} \left[\hat{z} C_{\bar{\gamma}} + C_N\right]^{-1} \bar{g}_N(\mathbf{x}, \theta, b) \end{aligned} \tag{8}$$

where $\hat{z}$ is computed using covariance matrix $C_{\gamma+N}$ from the noisy image. Thus, to estimate the original signal, 19 noise covariance matrices C_N must be estimated, 1 per subband, since $C_{\gamma+N} = C_\gamma + C_N$.

In the case of additive white noise, the noise covariance matrix is simply

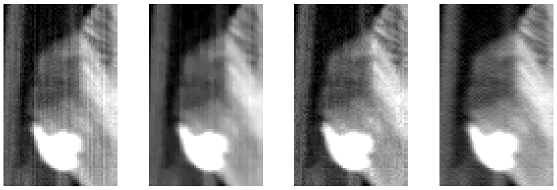
$$C_{N_w} = \sigma_{N_w}^2 \mathcal{I} \tag{9}$$

where $\sigma_{N_w}^2$ is the variance of the additive noise and $\mathcal{I}$ is the 15x15 (or 14x14) identity matrix. To extend GSM denoising to the spectral NU noise model for vertical noise, we must fully characterize the 19 covariance matrices, $C_{NU}(\sigma_V, \theta, b)$, where σ_V is the standard deviation of the additive vertical noise under the spectral model.

To analyze the behavior of the $C_{NU}(\sigma_V, \theta, b)$, a random sample of 20 LWIR images were selected and deformed over the range $\sigma_V \in \{0.0001, 0.0003, ..., 0.1\}$ to obtain the structure of the C_{NU} matrix parameterized on (σ_V, θ, b). Each C_{NU} was directly computed using the formula

$$C_{NU} = C_{NU+\gamma} - C_\gamma \tag{10}$$

where C_γ was obtained from the pristine (non-distorted) image and $C_{NU+\gamma}$ was obtained from the distorted image.



(a) Original (b) TVNUC (c) MIRE (d) GSMPP

Fig. 4. Algorithm performance on image with real vertical non-uniformity.

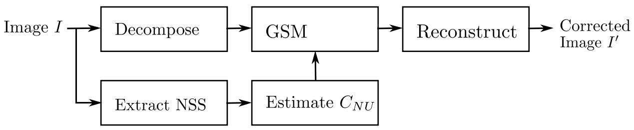


Fig. 5. Main steps followed to perform denoising process for GSMPP.

These samples of C_{NU} were appropriately averaged to estimate a content-independent noise covariance. To account for differences in power observed at the covariance matrices per each image, the resulting estimate of C_{NU} is normalized by dividing out the absolute sum of the covariance matrix coefficients. Plots of the individual terms in C_{NU} for increasing σ_V are depicted in Fig. 3. Notice that with increasing σ_V, each coefficient in each C_{NU} matrix is monotonic, and thus there is a one-to-one correspondence between each covariance matrix and the parameter σ_V.

To model this correspondence, polynomial fits were applied. Figs. 3(a), 3(b), 3(c), and 3(d) demonstrate that the 6th degree polynomial fit to these sampled covariance coefficients is adequate for modeling the monotonic behavior.

Given the functional mapping for C_{NU} at each subband, denoising then relies on an accurate estimate of σ_V, estimated using PNU as developed in section III. Given these components, the denoising algorithm GSMPP is complete, as depicted fully in Fig. 5. The model first extracts the NSS features associated with the PNU metric from an input image. The PNU metric returns an estimate of σ_V from which noise covariance matrices are generated. Each of these matrices is multiplied by the sum of the absolute values in the corresponding input image matrices, to account for the total image power. The input image is then split into steerable pyramid subbands. The GSM denoising algorithm uses the generated $C_{NU}(\sigma_V, \theta, b)$ followed by image reconstruction.

V. Performance of NUC Models

For evaluation, full-reference measurements such as the mean-squared error (MSE), peak signal-to-noise ratio (PSNR), and the Structural Similarity Index (SSIM) were used. MSE measures the average squared deviation, PSNR measures the error relative to the maximum signal level, and SSIM provides a perceptually relevant measure of quality.

The input image corpus was distorted using the spectral noise model using σ_V randomly sampled from $[0.0025, 0.025]$. This range was chosen subjectively to provide perceptually separable samples of non-uniformity, ranging from barely visible to almost overpowering. Results from each algorithm were computed by comparing the noise corrected image with the pristine image. Table I depicts mean measurements, and Fig. 6 depicts boxplots using SSIM.

TABLE I
Mean full-reference measurements of vertical non-uniformity correction performance over 468 images

Model	MSE	PSNR	SSIM
NONE	0.018	40.2	0.853
TVNUC	0.017	41.0	0.879
MIRE	0.034	33.7	0.874
GSMPP full	**0.012**	**44.2**	**0.956**
GSMPP w/σ_V	0.011	45.1	0.970
GSMPP w/C_{NU}	0.010	46.4	0.978

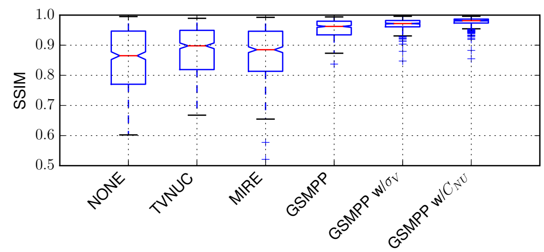


Fig. 6. SSIM measurements per image for each algorithm.

In both Fig. 6 and Table I, "NONE" compares the noisy images directly to the pristine images, to provide a baseline for comparison. MIRE appears to increase the MSE and PSNR, a result of the model's flicker-based noise assumption. TVNUC appears to produce an improvement relative to MIRE, since reducing strong localized diagonals effectively reduces non-uniformity. The "GSMPP full" model is a full model as depicted in Fig. 5. The "GSMPP w/σ_V" model uses the ground truth σ_V in place of PNU. The "GSMPP w/C_{NU}" model uses the noise covariance matrix computed using both reference and distorted images. The proposed GSMPP model outperforms the other models considered here. Given these results, improving PNU and refining estimated C_{NU} matrices can greatly improve GSMPP.

As a final comparison, an image with a vertical striping pattern was selected from MORRIS. Fig. 4 depicts the results of 3 correction algorithms on this image. TVNUC reduces the NU in the image, but visibly degrades image quality. Both MIRE and GSMPP remove most of the striping, with GSMPP producing the smoother result.

VI. Conclusion and Future Work

An algorithm has been developed and analyzed that can outperform state-of-the-art single image denoising algorithms when the noise is defined by spectral additive vertical striping. Since this algorithm requires a well-defined correlation structure of the noise, it is not necessarily specific to NU and can likely be generalized to other additive noise source patterns. Future work will investigate the computational complexity, corrective performance limitations, and noise characterization limitations of the GSMPP algorithm.

REFERENCES

[1] F. K. Amon, D. Leber, and N. Paulter, "Objective evaluation of imager performance," *Fifth International Conference on Sensing Technology*, pp. 47–52, 2011.

[2] J. M. Mooney, F. D. Sheppard, W. S. Ewing, J. E. Ewing, and J. Silverman, "Responsivity nonuniformity limited performance of infrared staring cameras," *Optical Engineering*, vol. 28, no. 11, pp. 281 151–281 151, 1989.

[3] J. M. Mooney, "Effect of spatial noise on the minimum resolvable temperature of a staring sensor," *Applied Optics*, vol. 30, no. 23, pp. 3324–3332, 1991.

[4] J. Lopez-Alonso, J. Alda, and E. Bernabeu, "Principal-component characterization of noise for infrared images," *Applied Optics*, vol. 41, no. 2, pp. 320–331, 2002.

[5] J. E. Pezoa and O. J. Medina, "Spectral model for fixed-pattern-noise in infrared focal-plane arrays," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Springer, 2011, pp. 55–63.

[6] F. Pérez, M. Nova, J. E. Pezoa, M. Figueroa, and S. N. Torres, "Spatial and frequency domain metrics for assessing fixed-pattern noise in infrared images," *SBMO/IEEE MTT-S International Microwave & Optoelectronics Conference*, pp. 1–5, 2013.

[7] F. Pérez, J. E. Pezoa, M. Figueroa, and S. N. Torres, "Empirical frequency domain model for fixed-pattern noise in infrared focal plane arrays," submitted to *Infrared Physics & Technology*. June 2014.

[8] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.

[9] R. Kafieh and H. Rabbani, "Wavelet-based medical infrared image noise reduction using local model for signal and noise," in *Statistical Signal Processing Workshop*, 2011, pp. 549–552.

[10] E. Vera, P. Meza, and S. Torres, "Total variation adaptive scene-based nonuniformity correction," in *Imaging Systems*. Optical Society of America, 2010.

[11] Y. Tendero, S. Landeau, and J. Gilles, "Non-uniformity correction of infrared images by midway equalization," *Submitted to Image Processing On Line*, pp. 1–23, 2012.

[12] T. Goodall, A. C. Bovik, and N. G. J. Paulter, "Tasking on natural statistics of infrared images," *IEEE Transactions on Image Processing*, unpublished.

[13] J. L. Rowe, *The Impact of Thermal Imaging Camera Display Quality on Fire Fighter Task Performance*. ProQuest, 2008.

[14] N. J. Morris, S. Avidan, W. Matusik, and H. Pfister, "Statistics of infrared images," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, 2007.

[15] G. C. Holst, *CCD Arrays, Cameras, and Displays*, JCD Publishing, Winter Park, Florida 32789 and SPIE, 1998.

[16] M. M. Hayat, S. N. Torres, E. Armstrong, S. C. Cain, and B. Yasuda, "Statistical algorithm for nonuniformity correction in focal-plane arrays," *Applied Optics*, vol. 38, no. 5, pp. 772–780, 1999.

[17] C. San Martina, S. N. Torresa, and J. E. Pezoa, "An effective reference-free performance metric for non-uniformity correction algorithms in infrared imaging system," in *Proc. of the 20th Annual Meeting of the Laser and Electro-optic Society*, Lake Buena Vista, Florida, USA, 2007, pp. 21–25.

[18] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Image Processing, International Conference on*, vol. 3. IEEE Computer Society, 1995, pp. 3444–3444.