Scene-based Non-uniformity Correction using Complementary Fixed Pattern Noise Models

Samet Almali Omer Faruk Adil
H. Seckin Demir
Microelectronics, Guidance and Electro-Optics Business Sector, ASELSAN Inc, Ankara, Turkey
sametalmali@aselsan.com.tr, ofadil@aselsan.com.tr, hsdemir@aselsan.com.tr

Abstract

We propose a novel scene-based non-uniformity correction (NUC) scheme for infrared focal plane array (FPA) detectors to account for both high-frequency and low-frequency fixed pattern noise (FPN). High-frequency FPN can be significantly reduced by the recent scene based NUC algorithms. However, low-frequency FPN caused by stray light, optical effects, heat dissipation and so forth, is commonly compensated by calibration based NUC methods. In this work, we aim to reduce both the low-frequency and high frequency components of FPN by using an efficient combination of registration based and constant statistics based approaches. We exploit scene variations through the video sequence and find the underlying low-frequency noise by smartly averaging frames based on their motion and detail content. Thus, we add the generalization power of constant statistics approach to existing scene-based NUC methods to obtain lower FPN in both high-frequency and low-frequency components. The performance of the proposed method is experimented on a public dataset corrupted by real FPN and evaluated by PSNR metric in comparison to a state-of-the-art scene-based NUC method.

1. Introduction

Infrared imagery has been of interest in biomedical, military, security, autonomous driving and various other fields. The focal plane arrays (FPA) used in infrared images have a well-known non-uniformity problem in which the detector pixels have non-uniform response to a given uniform radiance. This problem is alleviated by non-uniformity correction algorithms which are evolved into two main categories: calibration-based methods [7, 13, 2, 6] and scene-based methods [11, 10, 5, 3, 9, 12, 15, 14, 1]. In general, calibration-based methods are simple and easy to implement, however, they demand the capture of known uniform surfaces which leads to interruption of the video flow. This is not very desirable especially in critical applications that do not tolerate any loss of sight of target objects. Another drawback of these methods is that they often require special and expensive tools to generate reference surfaces to accomplish the calibration based NUC process. Such limitations gave rise to scene-based approaches in which the video sequence is not interrupted and only the scene image frames are used to estimate the FPN. Possible problems with these methods are the complexity of the algorithms and ghosting effects due to erroneous FPN estimations. However, recent methods [14, 1] are capable of producing both satisfactory FPN correction and negligible or even unnoticeable ghosting artifacts.

The history of scene-based NUC methods goes back to the study [11] that utilizes neural networks to find FPN correction terms based on constant statistics approach. Basically, the method assumes that the detector elements will eventually reach to a uniform temporal average if the scene content seen by the detector varies sufficiently. This assumption is actually a strong assumption to hold in finite image sequences as each pixel is not necessarily exposed to similar scene content variation. The violation of constant statistics assumption results in ghosting artifacts which causes unnatural images. The ghosting artifacts become severe as the high frequency components of the scene images increase. Although some other methods [10] try to improve on this approach, the ghosting artifacts still remain as a big problem. More recent studies [4, 16, 14] adopting LMS minimization approach for the errors found using registered image frames are more successful in reducing the ghosting artifacts. The registration of frames provides the compensation of scene content and significantly reduces the scene patterns burned into the non-uniformity correction terms. In [1], an adaptive strategy is employed in which correction terms are selectively updated considering registration errors and global motion requirements. Also, the method considers temporal consistency of the estimation steps via a Gaussian Mixture Model (GMM) structure to suppress spurious errors in the FPN estimation. These mechanisms are

978-1-5386-9294-3/18/$31.00 ©2018 IEEE

reported to result in very low ghosting artifacts, however, the method is still restricted to estimate high-frequency FPN and cannot compensate for the slow variations of FPN due to the small range of registration shifts. FPN with slower spatial variation than the registration shift amount cannot be detected and reduced. The residual low-frequency FPN has another side effect that degrades the contrast enhancement methods that are based on histogram modification. This noise causes an intensity variation that does not belong to the scene and consumes some of the dynamic range of the histogram. This false extension of the histogram dynamic range causes the histogram modification algorithm to stretch the input histogram less than the uncorrupted case which results in a loss of contrast gain.

The main contribution of our study is the utilization of a separate complementary estimation process that aims to find the low-frequency FPN in addition to the high-frequency FPN that is successfully handled by the existing methods. For this purpose, we employ a constant statistics approach. Although this approach was not very successful in high-frequency FPN reduction, it is a proper candidate for low-frequency FPN as it does not require a very accurate estimation. The performance of this method depends on the validity of the constant statistics assumption. Hence, we use different weights in the averaging process based on the assumptions of constant statistics: 1) The scene content varies significantly and 2) FPN is easier to discriminate when image contains less scene information. To fulfill the scene variation assumption, we use the inter-frame shift magnitude and limit averaging if the shift between the frames is not sufficient. We also adaptively increase the averaging weight if the image contains relatively less detail content. The edges of the image content are calculated and the average magnitude of the edges is used as the indicator of detail content of the image. By selecting the frames with sufficient translation and inversely weighting the selected frames with respect to their information content, we obtain a mean frame that models the underlying low-frequency FPN noise. This low-frequency FPN estimation is combined with the high-frequency FPN estimation and subtracted from the noisy image to obtain the corrected image.

The organization of the paper is outlined as follows. In Section 2, the method is explained in detailed sub-blocks. In Section 3, the experimentation process is described including the evaluation metrics, the public dataset used in experiments and the quantitative comparison of the proposed method and the scene-based NUC method presented in [1]. In Section 4, the concluding remarks are given.

2. Proposed Method

We propose a complementary low-frequency FPN estimation algorithm based on an improved constant statistics approach to combine with the state-of-the-art scene-based NUC method (IRSUM) described in [1]. The overview of the proposed method is depicted in the block diagram given in Fig. 1.

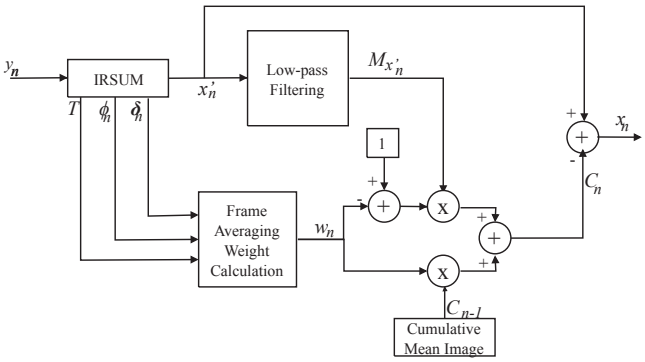


Figure 1. Block diagram of the proposed method. The further details of the "IRSUM" block is available in [1]

In the following sections, the details of each step in the proposed method are explained.

2.1. High-frequency Fixed Pattern Noise Estimation

As the first part of our method, the scene-based NUC method presented in [1] is employed for the high-frequency FPN estimation. In this method, well-known linear model is used for modeling the response of a detector as in (1).

$$y_n(i,j) = x'_n(i,j) \cdot a_n(i,j) + b_n(i,j) \tag{1}$$

where x'_n is the true response for the n^{th} frame, y_n is the associated detector output signal, (i,j) is the pixel coordinates and a and b are the gain and offset coefficients.

Considering this model, desired response may also be written as a function of the detector output as shown in Equation (2).

$$x'_n(i,j) = y_n(i,j) \cdot g_n(i,j) + o_n(i,j) \tag{2}$$

where backward gain $g_n(i,j)$ is equal to $1/a_n(i,j)$ and backward offset $o_n(i,j)$ is equal to $-b_n(i,j)/a_n(i,j)$.

According to this model, the observed signal is corrected by using the backward gain and offset coefficients. These coefficients can be estimated using different techniques. The employed method performs inter-frame registration and uses least mean squares (LMS) minimization for estimating the gain and offset coefficients.

For this purpose, IRSUM [1] method firstly registers the consecutive frames using an edge map. By this way, the amount of scene details can also be estimated using the high-frequency content of the frames. At the end of this registration operation, the scene detail magnitude ϕ_n and

shift amounts in both directions $\delta_{n\in\{x,y\}}$ are obtained. If the scene detail amount is higher than a certain threshold T, then the correction coefficients $a_n(i,j)$ and $b_n(i,j)$ are updated as given in (3).

$$\begin{aligned} a_{n+1}(p,q) &= a_n(p,q) - \eta_1 \cdot e_n(p,q) \cdot y_n(p,q) \cdot \xi_{st_n}(p,q) \\ b_{n+1}(p,q) &= b_n(p,q) - \eta_2 \cdot e_n(p,q) \cdot \xi_{st_n}(p,q) \end{aligned} \tag{3}$$

where, subscripts n and $n+1$ represents the current and next frame indices, (p,q) represents the index range within the registered overlap region between the previous and current image, η_1 and η_2 are the learn rates for the gain and offset updates, e_n is the error image between the corrected current frame and the previous frame, y_n is the current input frame and ξ_{st_n} is the spatio-temporal update mask which determines pixel locations at which the observation model will be updated.

Further details of the method are explained in [1].

2.2. Low-frequency Fixed Pattern Noise Estimation

In the second part of our algorithm, low-frequency FPN estimation process is applied. Basically, we aim to improve the high-frequency FPN estimation by a complementary low-frequency FPN estimation. This will not only provide better FPN reduction but also improve the contrast enhancement algorithm performance that commonly uses the output of NUC process.

2.2.1 Low-pass Filtering

When a new image frame arrives and gets processed through the high-frequency FPN estimation and correction steps, it becomes a candidate to contribute to the cumulative mean image which will then form the low-frequency FPN estimation. As we are seeking to capture the slowly varying components of the FPN, we discard the high-frequency component of the image frame by applying a Gaussian low-pass filter with a large kernel and standard deviation in order to obtain a heavily smoothed image as given in (4).

$$M_{x'_n} = x'_n * h \tag{4}$$

where $M_{x'_n}$ is the heavily smoothed image frame and h is the Gaussian low-pass filter kernel.

2.2.2 Moving Frame Selection

In order to improve the performance of constant statistics approach, we aim to use the frames that obey the assumptions of constant statistics. For this purpose, we eliminate the stationary frames in order not to violate the scene variation assumption. Thus, we obtain the motion information of an image frame by making use of the shift (δ_n) calculation made in the inter-frame registration step of IRSUM method. As a result, we selectively use the frames that are in motion with respect to the previous frame in order to create scene variation.

2.2.3 Frame Averaging Weight Calculation

For a further improvement in the constant statistics assumption validity, we make an adaptive weighting for the frame averaging based on the high-frequency information content within the image. Calibration based methods are able to effortlessly find the estimate of both the low-frequency FPN by capturing a known uniform surface so that the only variation in the detected image would belong to the FPN. This intuition is incorporated to our scene-based NUC by making a frequency content assessment of the image frame and assign more weight to more smooth images because they convey a more refined information about the underlying low frequency FPN. The high-frequency content in the image is assumed to be measured by the scene detail magnitude (ϕ_n) calculated in IRSUM algorithm. Then we use this value to calculate the adaptive averaging weight for the current frame as given in (5)

$$w_n = \begin{cases} \frac{\phi_n}{\phi_n + (T/c)} & \text{if } |\delta_n| > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where, w_n is the averaging weight for the cumulative mean image, ϕ_n is the scene detail magnitude value found in IRSUM algorithm, $|\delta_n|$ is the inter-frame shift magnitude, T is the detail magnitude threshold used in IRSUM algorithm to determine whether a frame contains sufficient detail information and c is a proportionality constant.

2.2.4 Cumulative Mean Image Update

After the frame averaging weight is calculated, the low-pass filtered image $M_{x'_n}$ will contribute to the new cumulative mean image proportional to its weight. This way, more informative frames about the underlying low-frequency FPN will be more effective in the resulting FPN estimation. This mechanism is shown in (6).

$$C_n = C_{n-1} * w_n + M_{x'_n} * (1 - w_n) \tag{6}$$

Where, C_n and C_{n-1} are the updated and previous cumulative mean images, $M_{x'_n}$ is the low-pass component of current image and w_n is the frame averaging weight calculated in (5). Due to constant statistics assumption, the cumulative mean image is expected to converge to the underlying FPN as the number or averaged frames increases. Thus, in practice, it is a good idea to allow for a sufficient number of averaged frames before applying the correction to the displayed image. The selective averaging strategy enables better estimation of low-frequency FPN at the expense

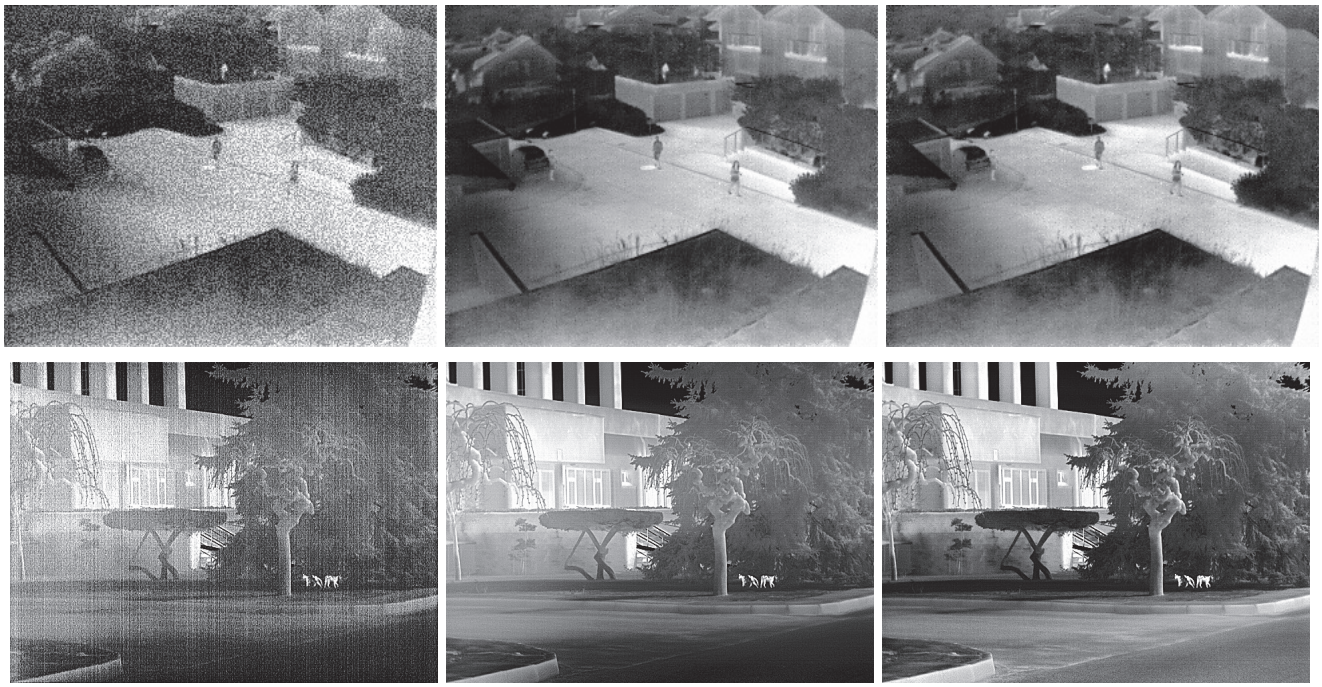


Figure 2. Demonstration of corrupted image, IRSUM result and our result. (Top: ETH dataset[8], bottom: our dataset)

of slower convergence of the cumulative mean image which eventually forms the low-frequency FPN model. The value range of w_n in (6) must be adjusted adequately to set a suitable trade-off between the convergence speed and undesired scene content sinking into the FPN model. In our method, w_n is determined adaptively and it has a bounded maximum value that reduces the chances of scene artifacts in the FPN model.

2.2.5 Low-frequency Fixed Pattern Noise Correction

At this step of the algorithm, we have an estimation of the low-frequency FPN. High-frequency FPN estimation and correction are applied as mentioned in Sec. 2.1 before the low-frequency FPN estimation and correction steps. Now low-frequency FPN correction is a simple subtraction of the estimated FPN as given in (7).

$$x_n = x'_n - (C_n - \overline{C_n}) \tag{7}$$

Where, x_n is the corrected output image, x'_n is the output of the high-frequency FPN correction (IRSUM), C_n and $\overline{C_n}$ are the cumulative mean image and mean value of cumulative mean image, respectively.

3. Experiments

The method in this work can be considered as a generic improvement step for any competitive scene-based NUC method. Thus, for the experimentations we qualitatively and quantitatively analyze the improvement over a state-of-the-art scene-based NUC method. We compared our method against IRSUM method [1] due to the fact that it is reported to have superior performance over other state-of-the-art methods. The experimentations are performed on the publicly available ETH-ASL Thermal Image Dataset [8] that are corrupted by a real high-frequency and low-frequency FPN measured from actual thermal imagers. The performance of the algorithms are compared by the Peak Signal to Noise Ratio (PSNR) metric explained in Sec. 3.1 that measures the similarity of the corrected images and the uncorrupted original images. Also, some sample result frames are displayed for qualitative analysis of the proposed method and the other method. The frames are taken both from the ETH-ASL [8] dataset and recordings from our infrared imagers.

3.1. Performance Metrics

The quantitative comparison of IRSUM and our improved method is performed by measuring the PSNR between corrected output images of both methods and the original clean images. PSNR yields how closely the corrupted images are corrected towards the original clean images. The PSNR measure of corrupted images are also calculated for reference. PSNR calculation is given in (8).

$$PSNR = 10log_{10}(peakval^2/MSE) \tag{8}$$

where, MSE is the mean-squared-error between the output and target images and $peakval$ is the maximum value that the input signal can take.

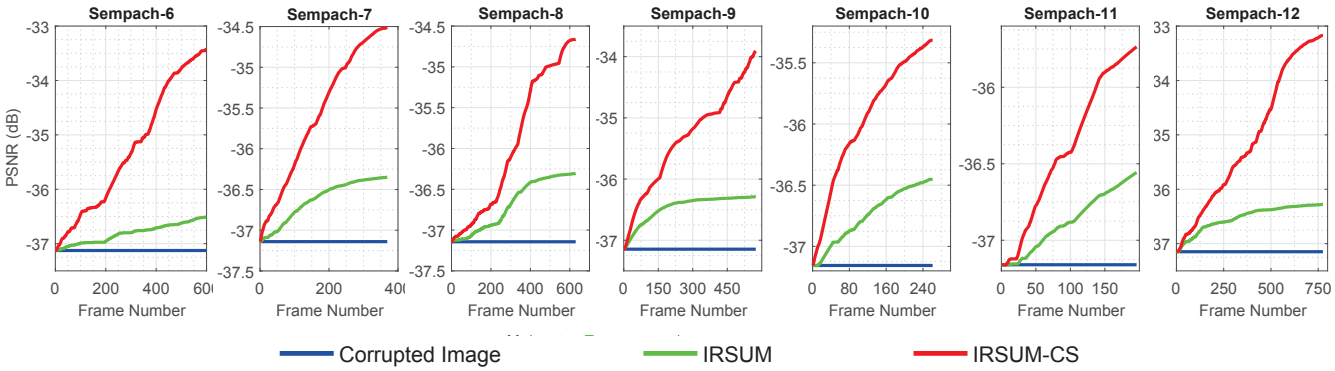


Figure 3. PSNR performance plots of the evaluated methods. The titles are given as in the original ETH-ASL Thermal Image Dataset [8]

3.2. Dataset

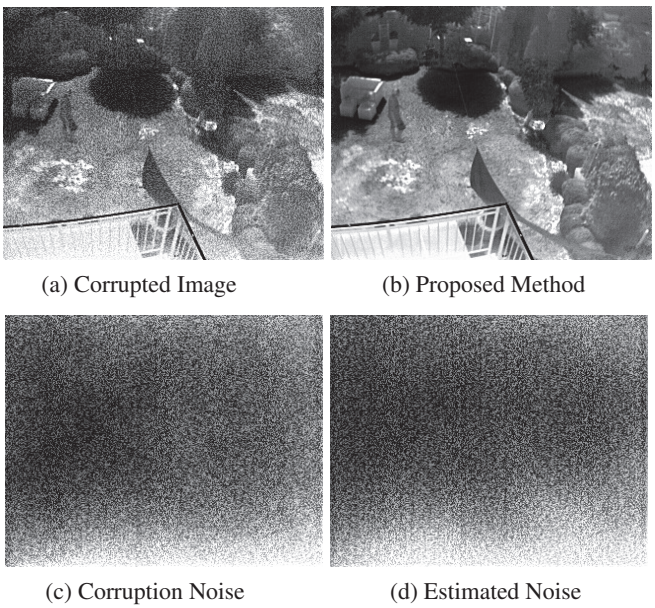
The quantitative evaluation of the proposed method is carried out on the ETH-ASL Thermal Image Dataset [8] which is publicly available[1]. This dataset features various scene contents such as roads, terrain, buildings, humans, animals, vehicles and so forth. Also, the sequence contains images both when camera is stationary and moving. Therefore, it is a suitable dataset to test scene-based NUC methods under the challenges of motion, stationary camera, deformable objects, non-rigid scene motion, changing illumination and so forth. The dataset is formed of images captured using a FLIR Tau 320 thermal imager which has the resolution of 324x256 and bit-depth of 16-bits. Dataset consists of 7 sequences with 3409 frames in total. The FPN is measured from an actual noisy camera image and added to the original clean images. Other image samples captured by our infrared imagers are used for visual comparison of our method. Those images have similar scene contents recorded in the resolution of 640x512 and bit-depth of 16-bits. This additional data is used to observe whether our method can achieve comparable performance on images from a completely different thermal imager and signal wavelength.

3.3. Results

In this section, we document the quantitative comparison of our method against the other state-of-the-art method based on PSNR metric using the uncorrupted images and correction results. Also, we demonstrate visual samples of output images to compare the correction performances of the methods.

The results of the experimentations are investigated using PSNR metric and sample output frame comparisons. During the execution of the algorithms, the parameters of the IRSUM algorithm and our method are kept fixed. We set scene detail threshold $T_\phi = 20$, registration quality threshold $T_\psi = 0.1$, gain learning rate $\eta_1 = 2 \cdot 10^{-9}$, offset learning rate $\eta_2 = 5 \cdot 10^{-3}$, spatial map outlier coefficient $c = 2$, GMM outlier coefficient $\lambda = 3$, GMM mean image learning rate $\alpha_1 = 0.1$, GMM variance image learning rate $\alpha_2 = 0.01$ for IRSUM and $T = T_\phi$, $c = 50$ and $\delta_c = 1$ for our low-frequency FPN correction algorithm.

In Figure 3, PSNR plots of the corrected images produced by the proposed IRSUM-CS and the IRSUM methods are shown. Also, the PSNR value of the corrupted input image is displayed for reference. It is noted that in all sequences of data, the proposed method performs better. The correction of the low-frequency FPN results in a significant effect in PSNR value due to the fact that this slowly varying FPN has a large magnitude. We also note that most of the PSNR curves of IRSUM-CS have not yet converged unlike the IRSUM PSNR curves which are observed to be closer to convergence. This is mainly because the constant statis-



(a) Corrupted Image (b) Proposed Method

(c) Corruption Noise (d) Estimated Noise

Figure 4. An example result of the proposed method on ETH dataset. a) Corrupted image constructed by adding synthetic noise to clean data. b) Denoised image result of the proposed method c) Noise pattern used for constructing the corrupted image d) Estimation of the noise pattern

[1] http://projects.asl.ethz.ch/datasets/doku.php?id=ir:iricra2014

tics approach could accept more frames to obtain a more accurate FPN model. We can expect to witness even better results with relatively longer image sequences.

Figure 2 shows some corrected output images of the proposed and compared methods. The left column images are corrupted images with real and synthetic FPN, the middle column images are the output of IRSUM method and the right column images are produced by the proposed IRSUM-CS method. On the top row, a sample frame from the ETH-ASL dataset [8] is shown. The input image is corrupted by both a real FPN with high- and low-frequency FPN components and a synthetic Gaussian noise. We note that the result of IRSUM method is quite impressive. But, when carefully viewed, some dark regions are spotted near the left and bottom regions. As a result, our method provides better detail perception especially near the house on top-left region and the part of the roof that resides on the bottom region of the image. On the bottom row, recordings from our thermal imager are shown. The left-side image displays the corrupted image with real measured FPN. Again, IRSUM performs rather well, however, the contrast and detail perception is more dramatic in this sample. The details of the trees are obviously stronger and the dark shade near the road on the bottom region is almost completely removed.

In order to observe the accuracy of the estimation of our FPN model, the added noise and the estimated noise are compared in Figure 4. The corresponding corrupted and corrected images are also displayed for reference. It is noted that not only the high-frequency Gaussian noise is successfully modeled, but also the slowly varying pattern is approximately estimated. This slow change pattern is difficult to detect for registration-based methods since these methods can only compare small neighborhood of pixel responses due to the fact that inter-frame shifts are commonly small. The constant statistics approach, on the other hand, exploit the whole image response and constitutes a complementary estimation to the state-of-the-art registration-based NUC methods.

4. Conclusions

In this work, we propose a novel scene-based NUC scheme that achieves superior performance by making use of complementary FPN models. Our method not only accounts for the high-frequency FPN as in the case of state-of-the-art methods, but also estimates the low-frequency component of the FPN by adopting a constant statistics model. The performance of the proposed method is evaluated and compared to a recent method both qualitatively and quantitatively on a publicly available dataset. The PSNR metric comparisons show that our method performs better in recovering the clean image. Also, visual comparisons demonstrate significant contrast and detail perception improvement.

References

[1] H. S. Demir and O. F. Adil. Irsum: Inter-frame registration based non-uniformity correction using spatio-temporal update mask. In *Applications of Digital Image Processing XLI*, volume 10752. International Society for Optics and Photonics, 2018.

[2] A. Friedenberg and I. Goldblatt. Nonuniformity two-point linear correction errors in infrared focal plane arrays. *Optical Engineering*, 37(4):1251–1254, 1998.

[3] R. C. Hardie, F. Baxley, B. Brys, and P. Hytla. Scene-based nonuniformity correction with reduced ghosting using a gated lms algorithm. *Optics express*, 17(17):14918–14933, 2009.

[4] R. C. Hardie, M. M. Hayat, E. Armstrong, and B. Yasuda. Scene-based nonuniformity correction with video sequences and registration. *Applied Optics*, 39(8):1241–1250, 2000.

[5] J. G. Harris and Y.-M. Chiang. Minimizing the ghosting artifact in scene-based nonuniformity correction. In *Proc. SPIE*, volume 3377, pages 106–113, 1998.

[6] S. Kim. Two-point correction and minimum filter-based nonuniformity correction for scan-based aerial infrared cameras. *Optical Engineering*, 51(10):106401–106401, 2012.

[7] D. L. Perry and E. L. Dereniak. Linear theory of nonuniformity correction in infrared staring sensors. *Optical Engineering*, 32(8):1854–1860, 1993.

[8] J. Portmann, S. Lynen, M. Chli, and R. Siegwart. People detection and tracking from aerial thermal views. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1794–1800, May 2014.

[9] B. M. Ratliff, M. M. Hayat, and J. S. Tyo. Generalized algebraic scene-based nonuniformity correction algorithm. *JOSA A*, 22(2):239–249, 2005.

[10] L. Rui, Y. Yin-Tang, Z. Duan, and L. Yue-Jin. Improved neural network based scene-adaptive nonuniformity correction method for infrared focal plane arrays. *Applied optics*, 47(24):4331–4335, 2008.

[11] D. A. Scribner, K. A. Sarkady, M. R. Kruer, J. T. Caulfield, J. Hunt, and C. Herman. Adaptive nonuniformity correction for ir focal plane arrays using neural networks. In *Proc. SPIE*, volume 1541, pages 100–109, 1991.

[12] S. N. Torres and M. M. Hayat. Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays. *JOSA A*, 20(3):470–480, 2003.

[13] E. Vera, P. Meza, and S. Torres. Total variation approach for adaptive nonuniformity correction in focal-plane arrays. *Optics letters*, 36(2):172–174, 2011.

[14] J. Zeng, X. Sui, and H. Gao. Adaptive image-registration-based nonuniformity correction algorithm with ghost artifacts eliminating for infrared focal plane arrays. *IEEE Photonics Journal*, 7(5):1–16, 2015.

[15] C. Zuo, Q. Chen, G. Gu, and W. Qian. New temporal high-pass filter nonuniformity correction based on bilateral filter. *Optical Review*, 18(2):197–202, 2011.

[16] C. Zuo, Q. Chen, G. Gu, and X. Sui. Scene-based nonuniformity correction algorithm based on interframe registration. *JOSA A*, 28(6):1164–1176, 2011.