Regularization approach to scene-based nonuniformity correction

Jun-Hyung Kim
Jieun Kim
Sohyun Kim
Joohyoung Lee
Boohwan Lee
Regularization approach to scene-based nonuniformity correction

Jun-Hyung Kim,* Jieun Kim, Sohyun Kim, Joohyoung Lee, and Boohwan Lee
Electro-Optic Directorate, 3rd R&D Institute, Agency for Defense Development, P.O. Box 35-31, Yuseong-gu, Daejeon 305-600, Republic of Korea

Abstract. Various scene-based nonuniformity correction (SBNUC) methods have been proposed to diminish the residual nonuniformity (RNU) of the infrared focal plane array (IRFPA) sensors. Most existing SBNUC techniques require a relatively large number of image frames to reduce the RNU. In some applications, however, there is not enough time for capturing a large number of image frames prior to the camera operation, or only several image frames are available to users. A new scene-based approach that can correct the RNU using only several image frames is proposed. The proposed method formulates the SBNUC process as an energy minimization problem. In the proposed energy function, we introduce regularization terms for the parameter regarding the responsivity of the IRFPA as well as for the true scene irradiance. Correction results are obtained by minimizing the energy function using a numerical technique. Experimental results demonstrate the effectiveness of the proposed method. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.OE.53.5.053105]

Keywords: residual nonuniformity; fixed-pattern noise; infrared focal plane array.

1 Introduction

The spatial noise caused by nonuniformity of individual detector elements in the infrared focal plane array (IRFPA) can limit the overall performance of imaging systems. In general, the responsivity of individual detector elements is assumed to be linear. Therefore, it is possible to correct the nonuniformity using the well-known two-point correction (TPC) method, where two blackbodies at different temperatures are employed.1 Since the nonuniformity normally drifts in time2 and the correction capability of the TPC is degraded with repeated operations,3 we need to recalibrate the IRFPA. In practice, there are two difficulties in using the TPC method to compensate the residual nonuniformity (RNU) resulting from the temporal drift or repeated operations: (1) we have to maintain two distinct heat sources in imaging systems; and (2) real-time video operation is interrupted during the correction process.4

Various scene-based nonuniformity correction (SBNUC) algorithms have been proposed to solve these problems. In general, SBNUC schemes can be broadly divided into two categories: constant statistics (CS) methods4−6 and least mean square (LMS) methods.7−10 The original CS method assumes that the temporal mean and standard deviation of each pixel are constant over time and space.5 The performance of the original CS method is reliable as long as the assumption is valid. As pointed out in Refs. 4 and 6, however, thousands of image frames are required to hold the assumption. Zhang and Zhao6 proposed a local constant statistics (LCS) method, which assumes that the temporal statistics are constant in a local region around each pixel. The LCS method improved the correction performance for the same number of input frames.4 Later, Zuo et al.5 generalized the LCS method by introducing a new constraint called multiscale CS.

In the LMS methods, the correction parameters are learned by minimizing the LMS error between corrected images and desired output images. The minimization is performed frame-by-frame using the stochastic gradient descent (SGD) technique.11 Since the desired output images are not available, spatially low-pass filtered input images are used as the desired ones.7 However, the performance of the LMS method is degraded at strong edge points as reported in Refs. 8 to 10. Thus, several methods improve the estimation accuracy of correction parameters by suppressing the influence of strong edge points in the minimization process. Vera and Torres8 adaptively adjust the learning rate which is a fixed parameter in the original SGD technique, according to the local spatial standard deviations of input images. In this way, the influence of strong edge regions, where the local standard deviation is large, is reduced in the minimization process. This method is further improved in Ref. 9 to remove burn-in artifacts caused by temporally slowly-varying image regions. An approach proposed in Ref. 10 updates correction parameters only when sufficient change occurs between consecutive images.

Lately, different error functions are proposed as alternatives to the LMS error in the original LMS method.12−14 Interframe registration-based LMS method12,13 first registers the previous corrected image with the current input image by assuming that only slight translational motion exists. Then, this technique minimizes the LMS error between the previously shifted image and the currently corrected image. Vera et al.14 minimize the total variation of corrected images to obtain the correction parameters.

Although these previous SBNUC algorithms may reduce the RNU, they have a common problem: a relatively large number of image frames are required to acquire the
Kim et al.: Regularization approach to scene-based nonuniformity correction

correction parameters.\textsuperscript{15} Two recently proposed SBNUC algorithms perform NUC using several image frames\textsuperscript{16} or even only two image frames.\textsuperscript{15} These methods can achieve good performance when the relative motion in successive image frames is a small translation along vertical or horizontal directions. However, large displacements between consecutive image frames can occur in some applications.\textsuperscript{17,18} To deal with the motion constraint of the previous approaches,\textsuperscript{15,16} we propose a new SBNUC method that estimates the correction parameters using several image frames. In the proposed method, we utilize the prior information on the parameters regarding the responsivity and the true scene irradiance. There is no restriction on the motion in successive image frames unless it is static.

The rest of this article is organized as follows. In Sec. 2, the proposed SBNUC method is detailed. In Sec. 3, the performance of the proposed method is evaluated. Conclusion is drawn in Sec. 4.

2 Proposed Method

In this section, we first formulate an optimization problem for correcting the RNU followed by its numerical solution.

2.1 Formulation

Let us assume that the characteristic of each detector element in the IRFPA is linear.\textsuperscript{5,7} Then, the acquired signal \( y(i, j, t) \) for the \((i, j)\)th detector element at time \( t \) is given by

\[
y(i, j, t) = x(i, j, t) + o(i, j, t),
\]

where \( o(i, j, t) \) represents the offset of each detector element and \( x(i, j, t) \) indicates the scene irradiance. Here, we assume that there is no gain nonuniformity since the offset component is the dominant source of the RNU.\textsuperscript{15,19,20} Given the image observation model [Eq. (1)], our objective is to estimate \( x(i, j, t) \) and \( o(i, j, t) \). This problem can be solved by minimizing the proposed energy function, which consists of three terms:

\[
E(x, o) = \sum_{n\in \Omega} \sum_{i\in I} D(x, o) + \lambda_x C_1(o) + \lambda_y C_2(x),
\]

where \( \Omega \) and \( I \) denote the set of given image frames and the image domain, respectively. \( \lambda_x \) and \( \lambda_y \) are regularization parameters for \( C_1(o) \) and \( C_2(x) \), respectively. The data-fidelity term \( D \) which measures the mismatch between the observed image and the estimates is given by

\[
D(x, o) = [y(i, j, t) - x(i, j, t) - o(i, j, t)]^2.
\]

Solving Eq. (3) alone is an underconstrained problem in which the number of unknowns is greater than that of equations. Thus, a regularization approach is taken to estimate \( x(i, j, t) \) and \( o(i, j, t) \) in this work. The regularization term for the offset \( o(i, j, t) \) is defined as follows:

\[
C_1(o) = \left[ \frac{\partial o(i, j, t)}{\partial t} \right]^2.
\]

\( C_1(o) \) is derived from the observation that the offset changes very slowly in time.\textsuperscript{3,4} In other words, the offset remains almost constant for several consecutive image frames.\textsuperscript{20} This regularization term favors the offset with small changes along the time axis. If we correctly estimate \( o(i, j, t) \) for the given image frames, their temporal variation is negligible, which means that \( C_1(o) \) is very close to zero.

The last term \( C_2(x) \) in Eq. (2) is introduced to regularize the scene irradiance \( x(i, j, t) \). In general, \( x(i, j, t) \) is smooth in the spatial domain. This fact is implicitly used in the original LMS method, where the desired image is the spatially low-pass filtered input image.\textsuperscript{7} Thus, it is natural for us to require \( x(i, j, t) \) to be smooth in the spatial domain. Since the degree of the spatial smoothness can be measured via the image gradient, \( C_2(x) \) is given by

\[
C_2(x) = w_x [y(i, j, t)] \left[ \frac{\partial x(i, j, t)}{\partial i} \right]^2 + w_y [y(i, j, t)] \left[ \frac{\partial x(i, j, t)}{\partial j} \right]^2.
\]

The smoothness term \( C_2(x) \) is proportional to the magnitude of the spatial intensity change of the scene irradiance. Therefore, the smoother the scene irradiance is, the smaller the value of \( C_2(x) \) is. However, the smoothness constraint is not appropriate at edge points as pointed out in Sec. 1. Since the spatial variation of \( x(i, j, t) \) is normally greater than that of RNU,\textsuperscript{2} the large spatial variation in the input image \( y(i, j, t) \) is mainly due to the edge points of \( x(i, j, t) \). We adaptively adjust the effect of the smoothness constraint according to the gradient of \( y(i, j, t) \). The weighting factors \( w_x[y(i, j, t)] \) and \( w_y[y(i, j, t)] \) in Eq. (5) are defined as follows

\[
w_x(y) = \left[ \left\| \frac{dy(i, j, t)}{d i} \right\| + \varepsilon \right]^{-1},
\]

\[
w_y(y) = \left[ \left\| \frac{dy(i, j, t)}{d j} \right\| + \varepsilon \right]^{-1}.
\]

The exponent \( \gamma \) controls the sensitivity to the spatial gradients of \( y(i, j, t) \) and \( \varepsilon \) is a small constant that prevents dividing by zero. Since the weighting factors are inversely proportional to the spatial gradients of \( y(i, j, t) \), the smoothness constraint has little effect on edge regions. These weighting factors are the same as smoothness weights for the image smoothing operator in Ref. 21.

2.2 Numerical Solution

The proposed energy function \( E(x, o) \) in Eq. (2), a function of two variables, is nonconvex. We minimize the energy function by solving two convex subproblems in an alternating way with initial estimates \( x^{(0)} = y \) and \( o^{(0)} = 0 \):

\[
x^{(n)} = \arg\min_x E[x, o^{(n-1)}],
\]

\[
o^{(n)} = \arg\min_o E[x^{(n)}, o].
\]
\( o^{(n)} = \arg\min_o E[x^{(n)}, o] \). \hspace{1cm} (9)

The above process is repeated until there is no significant change in the estimates \( x^{(n)} \) and \( o^{(n)} \).

We compute the partial derivative of Eq. (2) with respect to \( x \) in order to solve Eq. (8). First, we represent Eq. (2) in matrix notation as follows:

\[
\sum_{i,j \in \Omega} [y(i,j) - x(i,j) - o(i,j)]^T [y(i,j) - x(i,j) - o(i,j)] + \lambda_i [x(t)^T D_i^T W_i(t) D_i x(t) + x(t)^T D_j^T W_j(t) D_j x(t)],
\]

where \( y(t) \), \( x(t) \), and \( o(t) \) are the lexicographically ordered vectors corresponding to the acquired signal, the scene irradiance, and the offset at time \( t \), respectively. \( W_i(t) \) and \( W_j(t) \) denote diagonal matrices containing values of the weighting factors \( w_i(y) \) and \( w_j(y) \) at time \( t \), respectively. \( D_i \) and \( D_j \) are the backward difference operators along \( i \) and \( j \) directions, respectively. Note that the regularization term for the offset \( C_1(o) \) is omitted in Eq. (10) since \( C_1(o) \) is constant with respect to \( o \). Differentiation with respect to \( o \) produces

\[
\frac{\partial E}{\partial o} = \sum_{i,j \in \Omega} [A'(i,j) o(i,j) - b'(i,j)],
\]

where

\[
A'(i,j) = I + \lambda_i D_i^T D_i, \quad b'(i,j) = y(i,j) - x(i,j).
\]

Therefore, we solve a large system of linear equations [i.e., \( A'(i,j) x(i,j) = b'(i,j) \)] for each \( x(t) \). The conjugate-gradient (CG) method is used to solve the linear equations in this work since the matrix \( A \) is sparse, symmetric, and positive definite.22

We rewrite Eq. (2) in different matrix notation from that represented in Eq. (10) to solve Eq. (9):

\[
\sum_{i,j \in \Omega} [y(i,j) - x(i,j) - o(i,j)]^T [y(i,j) - x(i,j) - o(i,j)] + \lambda_o [o(i,j)^T D_o^T D_o o(i,j)],
\]

where \( y(i,j) \), \( x(i,j) \), and \( o(i,j) \) are vectors formed by lexicographically stacking the acquired signal, the scene irradiance, and the offset in the time domain, respectively, for each detector element located in \( (i,j) \). \( D_i \) is the temporal backward difference operator. Similar to Eq. (10), we exclude the regularization term \( C_2(x) \) in Eq. (14) since it is constant with respect to \( o \). Differentiation with respect to \( o \) produces

\[
\frac{\partial E}{\partial o} = \sum_{i,j \in \Omega} [A' o(i,j) - b'(i,j)],
\]

where

\[
A' = I + \lambda_o D_i^T D_i.
\]

\( I \) denotes identity matrix in Eq. (16). The CG method is used here again to obtain the values of the offset for each \( o(i,j) \).

### 3 Simulation Results

To our best knowledge, no study has been reported on correcting the RNU with several image frames that have large displacements. Therefore, no comparison is made with existing SBNUC methods in this work. The regularization parameters are empirically set to \( \lambda_o = 6 \), \( \lambda_i = 0.1 \) for all experiments in this article, and the value of \( \gamma \) is determined

---

**Fig. 1** Simulated nonuniformity images.
to be 0.2. At first, we test the convergence of the proposed method with eight synthetic images as shown in Fig. 1. The eight images are generated by adding the artificial offset\textsuperscript{9,15} to calibrated infrared images captured by a 320 × 256 InSb focal plane array camera operating in the 3 to 5 μm range. The RNU is generally composed of two patterns, the low-frequency one and white noise-like one as reported in Ref. 23. However, only the white noise-like pattern is usually prominent to observers together with natural scenes. This is due to the masking effect of the human visual system, which attenuates contrast sensitivity at low-spatial frequencies.\textsuperscript{24} Therefore, the artificial offset is generated as realizations of independent identically distributed Gaussian random variables.\textsuperscript{9}

We plot the proposed energy function in Eq. (2) against the number of iterations. As shown in Fig. 2, the value of the energy function drops quickly. We obtain good results with 11 iterations in our experiment. In Fig. 3, we present images corrected by the proposed NUC method. Close-up views of some parts of the images are depicted in Fig. 3(b) to help the reader observe the visual quality improvement. The proposed

<table>
<thead>
<tr>
<th>Image #</th>
<th>Input 2 images</th>
<th>4 images</th>
<th>6 images</th>
<th>8 images</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.65</td>
<td>29.61</td>
<td>31.4</td>
<td>32.19</td>
</tr>
<tr>
<td>2</td>
<td>27.64</td>
<td>29.52</td>
<td>31.37</td>
<td>32.17</td>
</tr>
<tr>
<td>3</td>
<td>27.62</td>
<td>—</td>
<td>31.34</td>
<td>32.19</td>
</tr>
<tr>
<td>4</td>
<td>27.59</td>
<td>—</td>
<td>31.27</td>
<td>32.17</td>
</tr>
<tr>
<td>5</td>
<td>27.59</td>
<td>—</td>
<td>—</td>
<td>32.12</td>
</tr>
<tr>
<td>6</td>
<td>27.57</td>
<td>—</td>
<td>—</td>
<td>32.01</td>
</tr>
<tr>
<td>7</td>
<td>27.54</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>27.52</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Average</td>
<td>27.59</td>
<td>29.57</td>
<td>31.35</td>
<td>32.14</td>
</tr>
</tbody>
</table>

Table 1: Peak signal-to-noise ratio results of the proposed method with various numbers of input images.
method produces acceptable results no matter how complex the spatial distribution is, as shown in Fig. 3.

We investigate the effect of the number of input images on the estimated scene irradiance using the eight synthetic images. Table 1 shows the peak signal-to-noise ratio results for different numbers of input images. As the number of input images increases, we can obtain more accurate results. This can be described by the fact that the information gained from consecutive image frames leads to high-quality NUC results and enhanced temporal consistency in the offset. Note that, however, raising the number of input images increases processing time. Thus, selecting a proper number of input images demands trade-off between the computational complexity and the image quality.

We also perform an experiment on two sets of real infrared images as shown in Fig. 4. We have collected the two sets of images using a 640 × 512 InSb focal plane array camera operating in the 3 to 5 μm range. One set of images in Fig. 4(a) shows drastic intensity changes among them due to atmospheric effects. The other set of images in Fig. 4(b) has relatively large motion. Objective results for the proposed method are provided in Table 2. We employ a roughness metric8,9,15 which is defined by

$$\rho = \frac{\| h_1 * I \|_1 + \| h_2 * I \|_1}{\| I \|_1},$$

where $h_1$ and $h_2$ are horizontal and vertical difference filters, respectively, $I$ represents the image under test, $\| \cdot \|_1$ is the $L_1$ norm, and $*$ denotes discrete convolution. The roughness metric $\rho$ measures the amount of high-frequency energy due to the RNU. As pointed out in Ref. 9, the metric cannot distinguish between true high-frequency energy and that from the RNU. However, the metric can be a useful indicator of the RNU to some degree when taken along with subjective evaluation.9

The correction results of the proposed method are depicted in Figs. 5 and 6. Similar to the simulated nonuniformity case, our method consistently suppresses the RNU as shown in Figs. 5(b) and 6(b). We also present the difference images between the input and the corrected in Fig. 7 to visualize the RNU corrected by the proposed method.

We have implemented the proposed method using C language. The simulation is performed on a PC with an Intel i7

<table>
<thead>
<tr>
<th>Image #</th>
<th>Fig. 4(a) Input</th>
<th>Fig. 4(a) Proposed</th>
<th>Fig. 4(b) Input</th>
<th>Fig. 4(b) Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2211</td>
<td>0.0717</td>
<td>0.2183</td>
<td>0.0626</td>
</tr>
<tr>
<td>2</td>
<td>0.1290</td>
<td>0.0798</td>
<td>0.3981</td>
<td>0.0618</td>
</tr>
<tr>
<td>3</td>
<td>0.4618</td>
<td>0.0648</td>
<td>0.4338</td>
<td>0.0465</td>
</tr>
<tr>
<td>4</td>
<td>0.7389</td>
<td>0.0484</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>0.3975</td>
<td>0.0685</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Fig. 4 Real infrared images.
Fig. 5 NUC results on images shown in Fig. 4(a). (a) Whole images. (b) Close-ups of the input (top) and the corrected (bottom).

Fig. 6 NUC results on images shown in Fig. 4(b). (a) Whole images. (b) Close-ups of the input (top) and the corrected (bottom).
3.40-GHz CPU and 4-GB memory. Our optimization procedure takes 5.6 and 3.8 s for the set of images in Figs. 4(a) and 4(b), respectively.

4 Conclusion
We presented a regularization approach to SBNUC with several image frames. Our work formulated the SBNUC process as the energy minimization problem that incorporates the slowly varying nature of the detector responsivity and the smoothness constraint for the scene irradiance. In the proposed method, no assumption was made about the motion among input images except that the motion is static. Therefore, the proposed method can be used in applications where only several image frames are available and large displacements exist among the given images. Simulation results on both synthetic and real infrared images demonstrated that the proposed method can reduce the RNU. In future works, we plan to apply more advanced numerical techniques to reduce the computational complexity of the proposed method.

References

Jun-Hyung Kim received his BS and PhD degrees in electronic engineering from Korea University in 2006 and 2012, respectively. He has worked for the Agency for Defense Development since 2012. His current research interests are in the area of image processing and infrared imaging system. He is a member of SPIE.

Jieun Kim received her BS degree in electrical engineering from Busan National University in 2002, and her MS degree in electrical engineering from KAIST in 2004. She has worked for the Agency for Defense Development since 2004. Her current interests include digital image processing, target detection, and tracking.

Sohyun Kim is currently a research member at the Agency for Defense Development in Korea, and has over 10 years of experience in developing electro-optic systems. Her experience includes target detection algorithm design and real-time implementation for video tracker for infrared images. She holds a BS in physics from Sogang University and an MS in information and communications from Gwangju Institute of Science and Technology.

Jooyoung Lee received his BS and MS degrees in electronics engineering from Dankuk University, Republic of Korea, in 1990 and 1992, respectively. Since 1992, he has been a principal researcher in the Electro-Optics Laboratory at the Agency for Defense Development. His research interests include analog and digital signal processing for IRST, low-noise electronics, IRST system test and evaluation for small target detection.

Bochwan Lee received his BS, MS, and PhD degrees in electrical engineering and computer science from Kyungpook National University, Daegu, Republic of Korea in 1991, 1993, and 2006, respectively. He has worked as a principal researcher for the Agency for Defense Development since 1993. His current interests include digital image processing, target detection, and tracking.