Optical Engineering

SPIEDigitalLibrary.org/oe

Scene-based nonuniformity correction method using multiscale constant statistics

Chao Zuo
Qian Chen
Guohua Gu
Xiubao Sui
Weixian Qian

Scene-based nonuniformity correction method using multiscale constant statistics

Chao Zuo
Qian Chen
Guohua Gu
Xiubao Sui
Weixian Qian
Nanjing University of Science and Technology
National Defense Key Laboratory of Optoelectronic Engineering
Nanjing 210094, China
E-mail: surpasszuo@163.com

Abstract. In scene-based nonuniformity correction (NUC) methods for infrared focal plane array cameras, the statistical approaches have been well studied because of their lower computational complexity. However, when the assumptions imposed by statistical algorithms are violated, their performance is poor. Moreover, many of these techniques, like the global constant statistics method, usually need tens of thousands of image frames to obtain a good NUC result. In this paper, we introduce a new statistical NUC method called the multiscale constant statistics (MSCS). The MSCS statically considers that the spatial scale of the temporal constant distribution expands over time. Under the assumption that the nonuniformity is distributed in a higher spatial frequency domain, the spatial range for gain and offset estimates gradually expands to guarantee fast compensation for nonuniformity. Furthermore, an exponential window and a tolerance interval for the acquired data are introduced to capture the drift in nonuniformity and eliminate the ghosting artifacts. The strength of the proposed method lies in its simplicity, low computational complexity, and its good trade-off between convergence rate and correction precision. The NUC ability of the proposed method is demonstrated by using infrared video sequences with both synthetic and real nonuniformity. [DOI: 10.1117/1.3610978]

Subject terms: infrared focal plane array; nonuniformity correction; multiscale constant statistics.

Paper 110255R received Mar. 16, 2011; revised manuscript received Jun. 6, 2011; accepted for publication Jun. 23, 2011; published online Aug. 4, 2011.

1 Introduction

Thermal array detectors, also known as infrared focal plane arrays (IRFPA), are a rapidly developing technology and are used in a wide range of industry, medical, and military applications. However, the nonuniformity in IRFPAs, which is due to pixel-to-pixel variation in the detectors' response, can considerably degrade the quality of IR images since it results in a fixed-pattern noise (FPN) that is superimposed on the true image.[1] Therefore, nonuniformity correction (NUC), being an indispensable key step, is applied to nearly all of the IRFPA-based engineering applications. Further, what makes the problem worse is that the nonuniformity varies over time and is closely related to external conditions,[2,3] which results in the failure of traditional reference-based NUC methods. In order to solve this problem, several scene-based nonuniformity correction (SBNUC) techniques have been recently developed.

There are two main categories of SBNUC: statistical methods[4–7] and registration-based methods.[8–10] Compared with registration-based methods, statistical approaches have been well studied because of their relatively lower computational complexity, smaller storage demands, and better real-time performance. The most well-known statistical method relies on the global constant statistic (GCS) assumption,[4,5] which states that the statistics of the observed scene become constant over time. This assumption requires that each detector in the array spend an equal amount of time observing a wide range of irradiance values. So it usually needs thousands of image frames in order to achieve a good NUC effect. When this assumption is violated, its performance is poor and ghosting artifacts are easily produced.[5,6] The local constant statistic (LCS)[7] method improves the GCS method and statistically assumes constant distribution in a local region around each pixel but uneven distribution in a larger scale. Thus, it improves the correction accuracy and reduces the ghosting artifacts that normally result in the GCS method. However, since no temporal window has been utilized in LCS, it fails to capture the drift of gain and offset coefficients. Besides, although LCS uses multiscale representation, the nonuniformity coefficients are actually decomposed into two components [low spatial frequency (LSF) and high spatial frequency (HSF)]. It has not fully utilized the multiscale scene information and a high-pass filtering function with proper filter size can achieve the same result.

In this paper, a new statistical NUC method called multiscale constant statistics (MSCS) is presented. We seek more effectiveness in the use of the scene information in different spatial scales. The MSCS assumption considers that it takes different frame samples for different spatial scales to meet the constant statistics constraint. The smaller the spatial scale is, the easier for the scale to meet the constant statistics assumption during a short period of time. So, in the MSCS method, the spatial range of gain and offset gradually estimates expands over time. Besides, we further improve the MSCS method by adding two extra features: an exponential window to capture the drift in nonuniformity and a tolerance interval for the acquired data to further reduce the ghosting artifacts. The final experimental results validate that the proposed method achieves a good NUC effect.

0091-3286/2011/$25.00 © 2011 SPIE

This paper is organized as follows. Section 2 reviews the GCS and the LCS methods. Section 3 presents the original form of the LCS method. Also, the MSCS method is introduced in detail. In Sec. 4, further enhancements on the MSCS method are discussed. The performance parameters are evaluated in Sec. 5. In Sec. 6, the proposed method is tested and compared with the GCS and LCS using sequences of infrared data with both simulated and real nonuniformity. Finally, we draw some conclusions of this paper in Sec. 7.

2 Global Constant Statistics and Local Constant Statistics

In this section, we first review the GCS method and its generalized form is given. Then the generalized form of the LCS is briefly recalled.

2.1 *Global Constant Statistics Method*

First, we assume that the photo-responses of an individual detector in a focal plane array can be characterized by a linear irradiance-voltage model[11]

$$y_{ij}(n) = g_{ij}(n) \cdot x_{ij}(n) + b_{ij}(n), \tag{1}$$

where n is frame index. The true scene value collected by the $(i, j)^{th}$ detector $x_{ij}(n)$ is scaled by a gain factor $g_{ij}(n)$ and shifted by an offset term $b_{ij}(n)$ to produce the observed detector output $y_{ij}(n)$. Nonuniformity correction is performed by applying a linear mapping to the observed pixel values $y_{ij}(n)$ to provide an estimate of the true scene value $x_{ij}(n)$ so that the detectors appear to be uniformly performing. For simplicity of notation, the pixel superscripts ij will be omitted and we use g and b to represent the gain and offset matrices of an IRFPA. Note that all matrix operations in this paper are preformed on a pixel-by-pixel basis.

The GCS method[4] assumes that the temporal means and mean deviations of input irradiances at each pixel are identical. For this assumption to hold, it is necessary that over time all possible scene irradiance values will be observed by all detectors. Then, $x(n)$ can be assumed to have zero mean and unity mean deviation without losing generality. The expressions are now written

$$\begin{cases} m = b \\ \sigma = g \end{cases}, \tag{2}$$

where m is the temporal mean and σ is the mean deviation of the observed signal. The estimated true signal is expressed as

$$x = \frac{y - m}{\sigma}. \tag{3}$$

Estimated mean and mean deviation of y can be calculated by the following recursive equations:

$$\begin{cases} m(n) = \dfrac{y(n) + (n-1)m(n-1)}{n} \\ \sigma(n) = \dfrac{|y(n) - m(n)| + (n-1)\sigma(n-1)}{n} \end{cases}. \tag{4}$$

Obviously, the offset obtained by Eq. (2) contains the direct current (DC) component and the gain is not normalized. The dynamic range of the corrected image may be inconsistent with the input, which is not what we want. So, the corrected images are usually rescaled to the same contrast and brightness as the original uncorrected image

$$\begin{aligned} x &= \frac{y - m}{\sigma} \times \langle \sigma \rangle + \langle m \rangle \\ &= \frac{y - \left(m - \frac{\sigma}{\langle \sigma \rangle} \times \langle m \rangle \right)}{\frac{\sigma}{\langle \sigma \rangle}} = \frac{y - \bar{b}_{cs}}{\bar{g}_{cs}}, \end{aligned} \tag{5}$$

where $\langle \bullet \rangle$ denotes the spatial average of all the elements in the array. Thus, we get the generalized form of the GCS. The normalized gain $\bar{g}_{cs}$ is $\frac{\sigma}{\langle \sigma \rangle}$ and the normalized offset $\bar{b}_{cs}$ is $m - \bar{g}_{cs}\langle m \rangle$.

2.2 *Local Constant Statistics Method*

The GCS method gives appropriate results only when all detector elements have been exposed to the same range of scene irradiance. In practical applications, this assumption is very difficult to be satisfied. Therefore, the LCS assumes that the temporal signal distribution is locally constant at each pixel but uneven in a larger scale. The gain and offset of the generalized form of the LCS (the LCS method presented in Ref. 7 is derived from the generalized form of the GCS, so we call it the generalized form of LCS to differentiate from the original form of LCS which given in Sec. 3.1) are calculated as follows:[7]

$$\begin{cases} \bar{g}_{lcs} = HSF(\bar{g}_{cs}) + I = HSF\left(\dfrac{\sigma}{\langle \sigma \rangle} \right) + I \\ \bar{b}_{lcs} = HSF(\bar{b}_{cs} | \bar{g}_{lcs}) = HSF(m - \bar{g}_{lcs}\langle m \rangle) \end{cases}, \tag{6}$$

where $HSF(\bullet)$ represents the high spatial frequency component of the input. I is an array of all ones. In the LCS, the HSF parts of the gain and offset image obtained by the GCS method are considered as the interference of the true scene signal. In order to remove the distribution variations of scene signal from the nonuniformity coefficients, the LCS method decomposes the gain image obtained by the GCS method into a multispectral and multiscale representation,[12,13] and removes its LSF part to get a better gain estimate first. Then, it applies the spectrum-shaping method to the offset image, removing its LSF part to obtain a better offset estimate. Note the gain $\bar{g}_{lcs}$ obtained by spectrum shaping is applied to normalize the offset image $\bar{b}_{cs}$ obtained by the GCS method. The corrected image of the LCS can be expressed as

$$x = \frac{y - \bar{b}_{lcs}}{\bar{g}_{lcs}}. \tag{7}$$

3 Multiscale Constant Statistics Method

3.1 *Original Form of Local Constant Statistic*

In Sec. 2.2, we have reviewed the correction formula of the generalized form of LCS. In order to better explain the LCS method, we covert the generalized form of LCS to its original form. Similar to Eq. (5), we perform inverse transform to Eq. (7)

$$\begin{aligned} x &= \frac{y - \bar{b}_{lcs}}{\bar{g}_{lcs}} = \frac{y - HSF\left\{ m - \left[HSF\left(\frac{\sigma}{\langle \sigma \rangle} \right) + I \right] \times \langle m \rangle \right\}}{HSF\left(\frac{\sigma}{\langle \sigma \rangle} \right) + I} \\ &= \frac{y - \left[HSF(m) - \frac{HSF(\sigma)}{\langle \sigma \rangle} \times \langle m \rangle \right]}{\frac{HSF(\sigma) + \langle \sigma \rangle}{\langle \sigma \rangle}} \end{aligned}$$

$$= \frac{y - \left[HSF(m) + \langle m\rangle - \frac{HSF(\sigma)+\langle\sigma\rangle}{\langle\sigma\rangle} \times \langle m\rangle\right]}{\frac{HSF(\sigma)+\langle\sigma\rangle}{\langle\sigma\rangle}} + \langle m\rangle$$

$$= \frac{y - [HSF(m) + \langle m\rangle]}{HSF(\sigma) + \langle\sigma\rangle} \langle HSF(\sigma) + \langle\sigma\rangle\rangle + \langle HSF(m) + \langle m\rangle\rangle. \quad (8)$$

The above derivation uses two simple relations: $HSF(HSF(\bullet)) = HSF(\bullet)$ and $\langle HSF(\bullet)\rangle = 0$. Referring to Eq. (5), we can get

$$\begin{cases} b_{lcs} = HSF(m) + \langle m\rangle \\ g_{lcs} = HSF(\sigma) + \langle\sigma\rangle \end{cases}. \quad (9)$$

From the gain and offset formula of the original form of the LCS, we can see that the LCS method uses the DC components in substitution for the LSF parts in the mean and mean deviation images. The spatial scale that meets the LCS constraint is equivalent to the spatial frequency of the high-pass filter (as mentioned above, in the LCS, a high-pass filtering function with proper filter size can achieve the same result as the spectrum shaping).

3.2 *Multiscale Constant Statistics*

The GCS algorithm assumes that the temporal mean and mean deviation of each pixel are constant over time and space. This assumption holds for typical video sequences where tens of thousands of frames are averaged. The LCS algorithm assumes that the temporal signal distribution is not constant at each pixel but is locally true. This assumption requires fewer frames since it is very likely that detectors within a small block around one pixel could be exposed to similar scenes. However, in the LCS, the definition of "local" is vague and determined by the characteristics of the noise pattern. It is unreasonable since the temporal signal distribution should be independent with the nonuniformity. When there is nonuniformity in the LSF domain, the LCS method has to increase the maximal pyramid level to make the local more "global." Under this condition, the LCS method is fairly similar to the GCS method and at least thousands of frames are needed to yield a good NUC result.

Before we introduce the MSCS constraint, we first make clear the relationship between the number of samples averaged and temporal signal distribution over space. We use a clean 5000-frame 14-bit infrared video sequence collected at 1 p.m. on a sunny day by using an 320×256 mid-wave IR (3 to 5 μm) cooled camera. The sequence was shot at the top of a tall building and the camera kept moving during the whole process. Some frames show half sky and half ground and some structures with strong irradiance repeatedly appear in the field of view. We have accumulated the temporal mean and mean deviation images and recorded them at 250, 500, 1000, 2000, and 5000 frames, respectively. Figure 1 shows the first frame of the test sequence and the temporal mean images at different stages. The spatial variance is used to measure the uniformity of the mean and the mean deviation images, and the variances of mean and mean deviation images at different stages are shown in Fig. 2.

It can be seen, with the increase in the number of averaged frames, the variances of the mean and the mean deviation images gradually decrease. Accordingly, the mean image shown in Fig. 1 becomes more and more uniform. This also means that we are getting closer and closer to the GCS constant. However, the final variance of the mean image is 1206 and the final variance of the mean deviation image is 146, which indicates that 5000 frames are not enough to meet the GCS constraint.

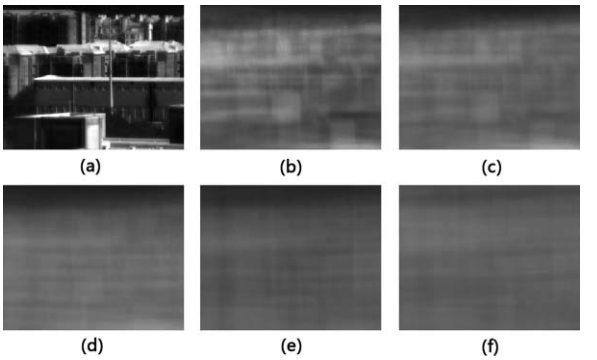


Fig. 1 Temporal mean images at different stages. (a) Frame 1 of the test video sequence taken at the top of a tall building. Temporal mean images of (b) 250, (c) 500, (d) 1000, (e) 2000, and (f) 5000 image frames. All images are scaled to the same display range.

To better analyze the temporal signal distribution in different spatial scales, we further decompose the temporal mean and the mean deviation images using the Laplacian pyramid (LP)[12,13]

$$\begin{cases} m = L_0(m) \oplus L_1(m) \oplus L_2(m) \oplus \ldots \oplus L_{N-1}(m) \oplus G_N(m) \\ \sigma = L_0(\sigma) \oplus L_1(\sigma) \oplus L_2(\sigma) \oplus \ldots \oplus L_{N-1}(\sigma) \oplus G_N(\sigma). \end{cases} \quad (10)$$

Thus, an original image is decomposed into levels of Laplacian images $L_0, L_1, \ldots, L_{N-1}$ and a Gaussian image G_N at the highest level of the pyramid (in this case, $N = 4$). We do not subsample the Gaussian image and apply a 9×9 normal filter at level 0 (the filter size doubles between two levels), thus, the decomposition and synthesis can be achieved by addition and subtraction between images. The spatial variances of the Laplacian images of the mean and mean deviation images at different stages are shown in Fig. 3.

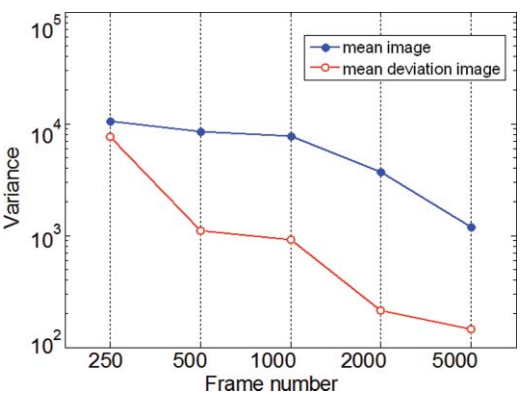


Fig. 2 Evolution of the variances of the temporal mean and the mean deviation images.

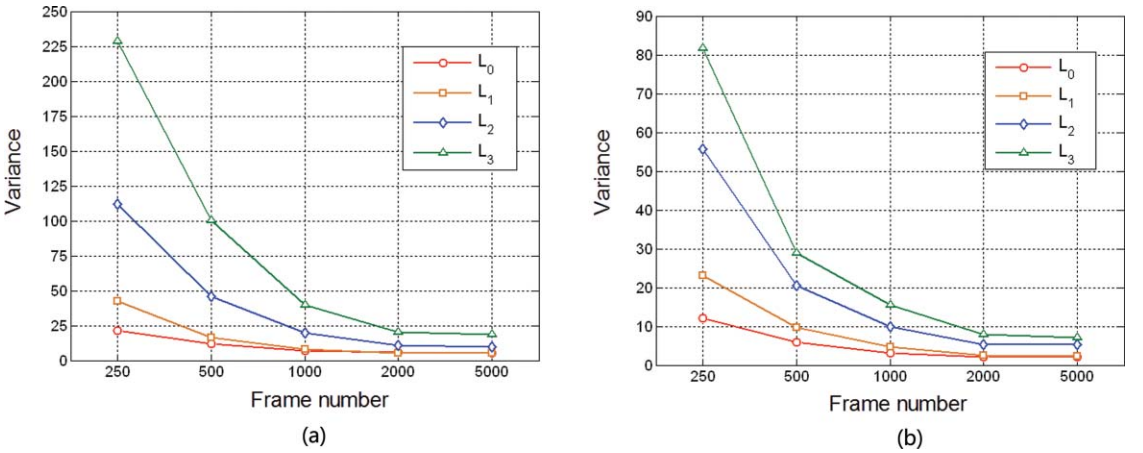


Fig. 3 Spatial variances of the Laplacian images of the (a) mean and (b) mean deviation images at different stages.

It can be seen from Fig. 3 that the variances of the HSF components are much smaller. Observing the temporal mean images in Fig. 1, we can also get the similar result: the uneven distribution of the temporal mean and the mean deviation images mostly concentrates on the LSF domain. This verifies the LCS constraint. In addition, it can also be found that the variances of HSF components dropped very fast, and the higher the spatial frequency, the lower the spatial variance. For the mean image, it only took $L_0(m)$ 250 frames to reach a variance of 25. However, to reach the same error, $L_1(m)$, $L_2(m)$, and $L_3(m)$ needed about 500, 1000, and 2000 frames, respectively. For the mean deviation image, we could get the similar relation. So we can draw the following conclusions: it takes different frame samples for different spatial scales to meet the constant statistics constraint. The smaller the spatial scale is, the easier it is for that scale to meet the constant statistics assumption during a short period of time, and vice versa. This is called the MSCS constraint.

A direct interpretation of the MSCS is that the closer the two detectors, the more likely they could "see" the same scene during a short period. For a single image, each pixel is isolated, so the constant statistics constraint is limited to the range of each individual pixel. As the number of averaged frames increases, the temporal signal distribution is assumed to be locally constant surrounding each pixel. As the number of frames further increases, the scale will gradually extend to the whole array. Meanwhile, the GCS assumption is satisfied.

Next, we will explain the MSCS assumption from another point of view. In the GCS algorithm, the mean and mean deviation are estimated with finite samples. The estimates of mean and standard deviation have the following relations with the true mean and standard deviation:[4]

$$MSE_m(n) = Var[m(n)] = E\{[m(n) - m]^2\} = \frac{\sigma_y^2}{n}, \quad (11)$$

$$MSE_{\sigma^2}(n) = Var[\sigma^2(n)] = E\{[\sigma^2(n) - \sigma^2]^2\} \approx \frac{2\sigma_y^4}{n}, \quad (12)$$

where MSE is the mean square error and $E[\bullet]$ and $Var[\bullet]$ represent the statistical expectation and statistical variance, respectively (since we have $E[m(n)] = m$ and $E[\sigma^2(n)] = \sigma^2$, here the mean square error is equivalent to the variance). It can be seen from the above equations that the errors of the mean and mean deviation estimates do not only decrease with the frame number, but are in proportion to the temporal variance of the observed data σ_y^2. We know that the mean and mean deviation images can be decomposed into a multiscale representation. The errors (variances) of the mean and the mean deviation estimates in one scale should be determined by the statistical characteristics of the observed data in that scale. This means that if the temporal variance of the signals at one level is smaller than that of others, fewer frames are needed to achieve the same estimation error.

Figure 4 shows a clean infrared image and its corresponding three Laplacian images and one Gaussian image using a four-level LP decomposition. The spatial variances for the corresponding Laplacian images are shown in Fig. 5. We can see that pixel-to-pixel correlations are largely removed in the Laplacian pyramid and the spatial variance of pixel values in the Laplacian pyramid is almost doubled between two levels.[13] Here, we extend the relation to temporal domain approximately

$$\sigma_{L_l(y)}^2 \approx 2^{l-k}\sigma_{L_k(y)}^2, \qquad \forall k, l \in Z, \quad (13)$$

where $\sigma_{L_l(y)}^2$ is the temporal variance of the observed data at level l of the LP and $\sigma_{L_k(y)}^2$ is the temporal variance of the observed data at level k of the LP. This relation has been verified with several simulations and experiment. It can be seen that, for a given detector, the standard deviation of its temporal values in smaller scales is much larger than those of the larger scales. From Eqs. (11) and (12), we know that in order to reach the same estimation error (variance), the numbers of frames needed for different scales are different. For the LP representation, the level increases by one, then to achieve the same error, nearly twice as many samples are required. This relation is in good accordance with the experimental results of Sec. 3.2. Besides, the derivations of Eqs. (11) and

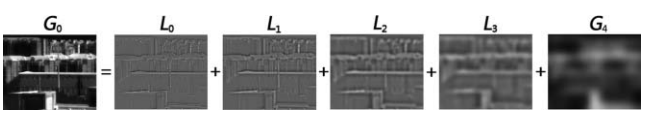


Fig. 4 Multiscale analysis of a clean infrared image: Laplacian pyramid parses an image into a series of Laplacian images of higher frequencies and a Gaussian image of lower frequency.

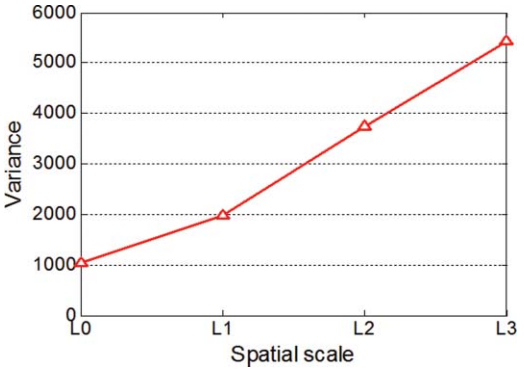


Fig. 5 Spatial variances of the Laplacian images of the corresponding infrared image shown in Fig. 4.

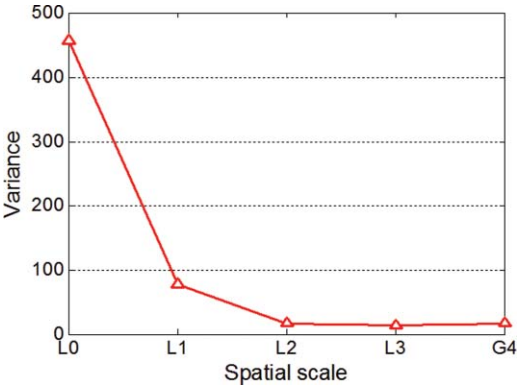


Fig. 7 Spatial variances of the images at different LP levels of the nonuniformity infrared image shown in Fig. 6.

(12) assume that pixel values are independent from frame to frame. Real-world images contain temporal correlations so that the variance of the mean and the mean standard deviation estimates may not always monotonically decrease as the frame number increases. But the pixel-to-pixel correlations are largely removed in the Laplacian images. So, it is reasonable to use Eqs. (11) and (12) for approximate error analysis.

3.3 *Nonuniformity Correction Using Multiscale Constant Statistics*

In this section, we discuss the NUC method based on the MSCS constraint. Figure 6 shows the multiscale representation of one typical nonuniformity image (a 40°C blackbody image captured by a one-point corrected uncooled infrared camera). Figure 7 shows the spatial variance of the images shown in Fig. 6.

Different from the scene image, the variance of the nonuniformity image is mainly distributed in the HSF domain. Most spatial variation is contained in the pyramid from level 0 to 1. This is exactly what we want, because only when one spatial scale meets the constant statistics constraint, can the nonuniformity of that scale be effectively corrected. As we pointed out earlier, the statistical constant constraint in smaller scales is easier to be satisfied. This means that we can quickly achieve a good uniformity correction result, without waiting until the GCS constraint is met when tens of thousands of frames have been used.

In addition, in stark contrast to the LSF image G_4 shown in Fig. 4, we can see the LSF part (G_4) of the nonuniformity image is very uniform and it has very little influence on the visual effect. This part can be ignored in NUC processing. Because, only when the error introduced by the uneven distribution of the true scene signal of one scale is less than the nonuniformity of that scale, can the NUC processing of that scale produce a positive effect. Usually, the error caused by a violation of the constant statistics constraint in larger scales is much greater than the residual nonuniformity in LSF.

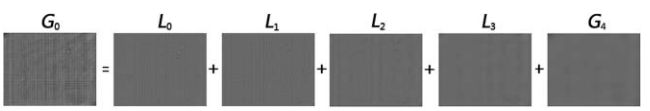


Fig. 6 Multiscale analysis of a nonuniformity image captured by an uncooled infrared camera.

Similarly, for the HSF components, if the frame samples are so few for one scale to meet constant statistics constraint, we should not use the statistics (mean and mean deviation) of that scale and assume the mean and mean deviation images in that scale are uniform images. In this way, the ghosting artifacts caused by the violation of the constant statistics constraint and the incorrect updating of NUC parameters can be greatly reduced.

Making the spatial scale of nonuniformity correction and spatial scale that meet the constant statistics constraint to synchronously expand is the basic idea of the MSCS-based NUC method. Figure 8 gives a schematic diagram of the MSCS, which can better illustrate this process. For the first frame, the constant statistics constraint is limited to the range of each individual pixel. So, we assume the mean and mean deviation are uniform images

$$\begin{cases} b = \langle m \rangle \\ g = \langle \sigma \rangle \end{cases}. \tag{14}$$

No NUC occurs at this time. As the number of accumulated frames increases, the spatial scale corresponding to level L_0 of the LP meets the constant statistics constraint. Then, we have

$$\begin{cases} b = L_0(m) \oplus \langle m \rangle \\ g = L_0(\sigma) \oplus \langle \sigma \rangle \end{cases}. \tag{15}$$

At this point, we only introduce the statistics of L_0 into the nonuniformity coefficients. Similarly, when the number of frames further increases, the scale extends to L_1level

$$\begin{cases} b = L_0(m) \oplus L_1(m) \oplus \langle m \rangle \\ g = L_0(\sigma) \oplus L_1(\sigma) \oplus \langle \sigma \rangle \end{cases}. \tag{16}$$

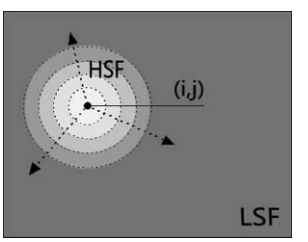


Fig. 8 Schematic diagram of the MSCS algorithm: the scope of MSCS gradually expands over time.

Followed by analogy, the spatial range of the MSCS expands to $L_2, L_3, \ldots, L_{N-1}$

$$\begin{cases} b = L_0(m) \oplus L_1(m) \oplus L_2(m) \oplus \ldots \oplus L_{N-1}(m) \oplus \langle m \rangle \\ g = L_0(\sigma) \oplus L_1(\sigma) \oplus L_2(\sigma) \oplus \ldots \oplus L_{N-1}(\sigma) \oplus \langle \sigma \rangle \end{cases}. \tag{17}$$

Then, most of the nonuniformity is contained in the scale that meets the constant statistics constraint and the residual nonuniformity in G_N can be ignored. The MSCS NUC reaches steady state and the formula for offset and gain are identical with those of the original form of the LCS.

The entire process described above expressed the basic idea of MSCS correction, i.e., the NUC parameters expand their sphere of influence in the spatial domain step by step. However, the effective radius of each Gaussian kernel is doubled between levels in the LP representation. It is clear the spatial range meets the constant statistics constraint should continuously expands rather than "jump" from one scale to another. From Eqs. (14)–(17), it can be found that the mean and mean deviation images are actually decomposed into LSF and HSF. Obviously, we can use a size-variable Gaussian filter to achieve the similar effect. Therefore, the offset and gain of the MSCS are calculated as follows:

$$\begin{cases} b(n) = HSF_{\mathrm{n}}[m(n)] + \langle m(n) \rangle \\ g(n) = HSF_{\mathrm{n}}[\sigma(n)] + \langle \sigma(n) \rangle \end{cases}, \tag{18}$$

where $HSF_{\mathrm{n}}(\bullet)$ denotes the HSF part of the mean and mean deviation images obtained by the size-variable Gaussian filter at frame n. Taking the mean image for example, its HSF component at frame n can be calculated as follows:

$$\begin{aligned} HSF_{\mathrm{n}}(m(n)) &= m(n) - LSF_{\mathrm{n}}(m(n)) \\ &= m(n) - m(n) \otimes G(\sigma_g(n)), \end{aligned} \tag{19}$$

where $G(\sigma_g(n))$ denotes the Gaussian kennel with a standard deviation parameter of $\sigma_g(n)$, $\otimes$ is the convolution operator. As discussed earlier, we know that if the spatial scale is doubled, the number of frames needed to meet the constant statistics constraint is approximately doubled accordingly. So, the relation between the number of averaged frame and the standard deviation parameter $\sigma_g(n)$ of the Gaussian filter could be approximated as linearity.

$$\sigma_g(n) = \begin{cases} \dfrac{(n-1)\sigma_{g\max}}{K} & n \le K \\ \sigma_{g\max} & n > K \end{cases}, \tag{20}$$

where $\sigma_{g\max}$ is the maximal standard deviation of the size-variable Gaussian filter. K represents the number of frames this maximal scale (corresponding to $\sigma_{g\max}$) needed to meet the constant statistics constraint. The spatial range of the MSCS gradually expands as n increases until K frames are used. The $(n-1)$ term is to ensure that $\sigma_g(n)$ is zero for the first frame, i.e., no NUC is performed.

The choice of $\sigma_{g\max}$ should take into account the spatial distribution characteristics of the nonuniformity, and it determines the value of K. K is chosen as $k\sigma_{g\max} + 1$, where k is a time constant parameter that determines how fast the spatial scale of MSCS expands. Pulsing one is to avoid dividing by zero under the special circumstance of $\sigma_{g\max} = 0$. The selection of $\sigma_{g\max}$ and k will be discussed in detail in subsequent sections.

4 Further Enhancements on the Multiscale Constant Statistics Method

In this section, we improve the MSCS method by adding two extra features: an exponential window for gain and offset estimation and a decision criterion to adaptively exclude abnormal data. These features can naturally help produce a faster update for the estimation and further reduce the ghosting artifacts.

4.1 *Multiscale Constant Statistics with an Exponential Window*

The recursive Eq. (4) consider the gain and the offset as stationary-unknown parameters. But, it was stated that the nonuniformity slowly and randomly changes over time. Therefore, the estimation provided by any NUC method must follow this drift. To better estimate changes of nonuniformity coefficients, here we introduce an exponential window[5] for gain and offset estimation. The exponential window emphasizes more on recent data, it provides the NUC method with the ability to follow changes in the operating point and drift in nonuniformity coefficients. By applying an exponential window function to the MSCS method, the recursive equations can be written as

$$\begin{cases} m(n) = \alpha(n)y(n) + [1 - \alpha(n)]m(n-1) \\ \sigma(n) = \alpha(n)|y(n) - m(n)| + [1 - \alpha(n)]\sigma(n-1) \end{cases}, \tag{21}$$

where $\alpha(n)$ is the time coefficient, which controls the window size and may change with n

$$\alpha(n) = \begin{cases} \dfrac{1}{M} & n \le M \\ \dfrac{1}{n} & N < n \le K \\ \dfrac{1}{K} & n > K \end{cases}. \tag{22}$$

The evolution of $\alpha(n)$ has three stages. The initialization stage includes the first M frames. This stage allows the gain and offset in small scale to quickly converge. Meanwhile, it avoids the large initial error that is normally introduced by using Eq. (4). Usually, M can be set as the number of frame needed for the scale of L_0 (9×9 Gaussian filter with the standard deviation of 1.5) to meet the constant statistics constraint. So, in our cases $M \approx 1.5k$ (since we have $\sigma_g(n) = \frac{(n-1)\sigma_{g\max}}{K}$ when $n \le k$, and $K = k\sigma_{g\max} + 1$, if $\sigma_g(M) = 1.5, M = \frac{1.5K}{\sigma_{g\max}} + 1 \approx 1.5k$). Followed by the initialization stage is the stage of MSCS scale expansion. The value of $\alpha(n)$ gradually reduces with the increase of n to allow more frames to be averaged. Finally, the algorithm reaches the steady state after K frames. The maximal scale is supposed to meet the constant statistics constraint and the nonuniformity has almost been removed. At that time, the algorithm should keep on iterating to guarantee the correction of the temporal drift of the FPN.

4.2 *Deghosting*

Ghosting artifacts are major problems of statistical SBNUC, which seriously affect the correction accuracy and visual quality. In the applications such as detection and tracking of targets, the problem is even worse because the slow speed

targets may be blended into the NUC coefficients and then leave an inverse ghost image in its original place. In the MSCS, ghosting artifacts are caused by two main factors: lack of global motion and extreme scene values. As for the former, the correction parameters are obtained by two temporal low-pass filters. If there is no image motion, it is implied that all signal energy is at zero frequency and the real scene will leak into the correction parameters. The ghosting caused by insufficient global motion is not uncommon among almost all SBNUC (Refs. 5, 7, and 16) and can be reduced by the change detection mechanism, which prohibits parameter updates when motion of the scene is insufficient.

As for the latter, our method generates running averages for the acquired data and their mean deviation for NUC parameters estimation. Though our method adopts a size-variable low-pass filter to better control the convergence of the gain and offset, it is evident that the running average is not a robust statistic since it is unduly affected by outliers or other small departures from model assumptions. The outlier values should be excluded or their contributions should be attenuated to prevent a large degree of influence on the parameters. Nevertheless, there is no rigid mathematical definition of what constitutes an outlier; determining whether or not an observation is an outlier is ultimately a subjective exercise. There are two kinds of methods available in the literature. In Refs. 14, 15, and 16, the outliers are defined based on experience, i.e., they are presumed to be of high local variance or around edges. The other method is based on statistics.[17,18] These methods usually make use of mean and variance of the previous data samples (or error samples) as a prior to examine the following data. Since in our method, the nonuniformity parameters are inherently calculated by mean and mean deviation statistics. We adopt a temporal statistics deghosting method similar to the method in Ref. 17. We use the temporal statistics of previous frames to decide whether the current pixel value will be used to compute the correction coefficients. Obviously, only when the previous statistics are reliable enough can we use them to determine whether a newcomer is outlier. So, during the initialization stage, no deghosting should be operated. After the initialization stage, we only utilize the statistics in the scales that are supposed to meet the constant statistics constraint. A tolerance interval for the acquired data is defined as follows:

$$[b(n-1) - Wg(n-1), b(n-1) + Wg(n-1)], \tag{23}$$

where the parameter W determines the width of the tolerance interval. The rule to operate deghosting is summarized as follows:

$$\begin{cases} \begin{cases} m(n) = \alpha(n)y(n) + [1-\alpha(n)]m(n-1) \\ \sigma(n) = \alpha(n)|y(n) - m(n)| + [1-\alpha(n)]\sigma(n-1) \end{cases} & \text{if } |y(n) - b(n-1)| \le Wg(n-1) \\ \begin{cases} m(n) = m(n-1) \\ \sigma(n) = \sigma(n-1) \end{cases} & \text{if } |y(n) - b(n-1)| > Wg(n-1) \end{cases}. \tag{24}$$

The updating of the correction coefficients is operated only when the observed signal were found reliable. This mechanism prevents biased estimates from abnormal updating caused by the outliers and helps to sample data better.

5 Parameters Evaluation

5.1 *Determining the Maximal Scale for Multiscale Constant Statistics*

The maximal standard deviation $\sigma_{g\max}$ of the size-variable Gaussian filter determines the largest scale of MSCS in steady state. If the largest scale of MSCS is equal to the image size, then the MSCS will gradually approach the GCS as the frame number increases. Obviously, we hope most of the noise patterns can be filtered out by the Guassian filter with a standard deviation of $\sigma_{g\max}$, and $\sigma_{g\max}$ should be as small as possible to avoid the unwanted errors introduced by the true scene signal. So, we can simply face the camera to blackbodies at different operation temperatures and get a set of images that contain only noise patterns. Then, we apply the size-variable Gaussian filter to these noise images. When the residual nonuniformity in the filter output of each image can be ignored as the standard deviation parameter increases. An ideal value of $\sigma_{g\max}$ is obtained. The $\sigma_{g\max}$ should generally not be too large since the residual FPN of IRFPAs mostly concentrates on the HSF domain after calibration by reference-based NUC methods.[7,19]

5.2 *Parameters Analysis*

There are two parameters can be used as control knobs in our algorithm: the time constant parameter k and the width of the tolerance interval W. We now study the impact of the two parameters on the noise compensation capability of our algorithm for the purpose of finding the best performing setting. To measure the overall NUC performance when the image is corrupted by both offset and gain nonuniformity, the peak signal-to-noise ratio (PSNR)[10,14] is used to quantify the differences between a clean reference image against its noisy and nonuniformity corrected versions, and it is defined as

$$PSNR(n) = 20\log_{10}\left[\frac{D}{RMSE(n)}\right], \tag{25}$$

$$RMSE(n) = \sqrt{\frac{\sum_{i,j}[x_{ij} - x_{ij}(n)]^2}{N}}, \tag{26}$$

where $x_{ij}(n)$ is the $(i,j)^{th}$ pixel's value of the true frame while $\hat{x}_{ij}(n)$ is the pixel's value of the corrected frame. N is the total number of pixels in the image. D represents the dynamic range of the images (we choose D as the average dynamic range of the input frames). The larger the PSNR value, the closer the estimated signals are to the true signals.

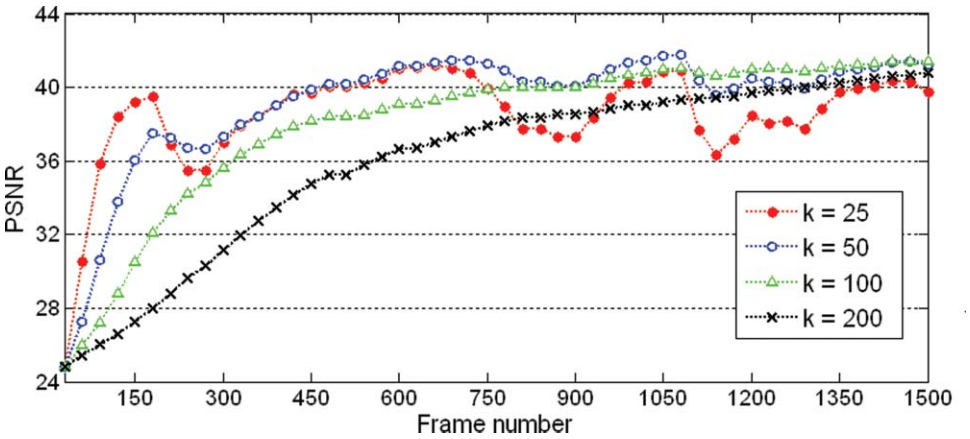


Fig. 9 PSNR versus frame number using different time constant parameters.

For convenience and in all simulations, the initial correction parameters have been initialized with unitary gain and null offset. The simulated gain and the offset are normally distributed and the nonuniformity is assumed to be distributed mainly within the scale corresponding to $\sigma_{g\,\max} = 5$.

5.2.1 *Time constant parameter analysis*

We use a clean 1500-frame outdoor infrared video sequence corrupted by simulated nonuniformity. To better study the convergence process with different time constants parameter k. The sequence presents a sufficient global motion and very little extreme scene, thus no deghosting is performed. We have studied four k values, respectively 25, 50, 100, and 200. The PSNRs of the corrected images versus frame number are shown in Fig. 9.

It can be seen that the speed of convergence decreases with the increase in the value of k. This is expected, since when k is small, the cutoff frequency of the temporal low-pass filter will be high and fewer frames will be averaged. Besides, if k is too small ($k = 25$), the curve becomes so volatile that after 200 frames the curve began to decline, which result from the lack of statistical gray-value diversity within less frames. This lack of diversity violates the MSCS assumption. Conversely, the steady-state error increases with the increase in the time constant. So, k can be set as 100, this value obtains a good trade-off between the speed and the steady-state accuracy of estimation. Furthermore, in real situations, the value of k needs to be adjusted according to the sample rate and the application field of the camera.

5.2.2 *Tolerance interval analysis*

The 1500-frame sequence with artificial nonuniformity used to test the deghosting performance has been taken using the same camera as in Sec. 3.2, corrupted by using synthetic gain and offset. The sequence typically presents complex background and extreme scene values such as strong edges of structures, high-gray-level objects caused by sunlight reflection, separation edges caused by horizon effects, or intensity transitions between sky and ground regions. To prevent ghosting artifacts, a proper W should be adopted to better classify the outliers and remove these anomalous observations that cause biased estimates. For a proper comparison, the time constants parameter k is chosen as 100 using the previous results. The PSNRs of the corrected images are shown in Fig. 10 by varying the tolerance interval W.

It can be seen that the PSNR is almost insensitive to the tolerance interval W for the first 200 frames. But after that, the deghosting plays a more important role in the convergence process. If a smaller value of W is used, the PSNR curve becomes rather smooth and inactive after 200 frames, which can be helpful to prevent ghosting. However, subsequent convergence of the algorithm becomes slower since only very few data points are used for estimations. On the other hand, a larger value of W tends to decrease the deghosting effect. It is recommended to choose $W = 2$ to improve the deghosting capabilities of our method. This choice yields best results of NUC with our default time constant parameter setting.

6 Nonuniformity Correction Performance

In this section, several tests on both simulated and real data are performed to verify the MSCS method and to compare the NUC performance of our method with the GCS and LCS.

6.1 *Applications to Simulated Infrared Data*

In this test, 1200-frame image sequence clipped from the same 5000-frame sequence as in Sec. 3.2 are artificially corrupted by FPN, and then the performance of the aforesaid three NUC methods will be quantitatively evaluated using the PSNR as a metric. The PSNR of the corrupted image sequences with simulated nonuniformity are about 24.2 dB for all of the frames. For this simulated nonuniformity, the maximum pyramid level in LCS is set as two for spectrum

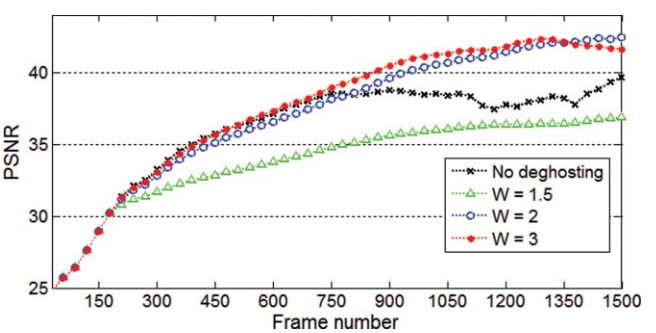


Fig. 10 PSNR results using a different W.

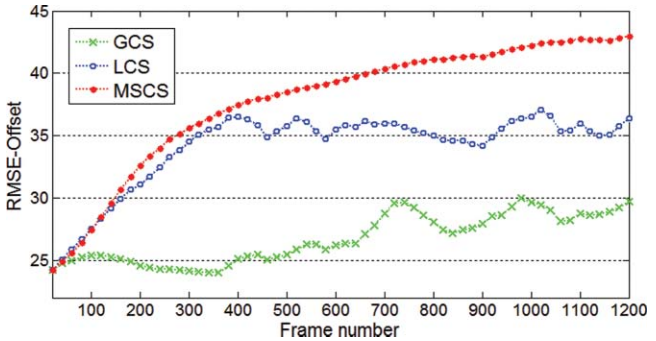


Fig. 11 PSNR results of the synthetic noisy test sequence corrected with different NUC methods.

shaping and the MSCS uses $\sigma_{g\,\max} = 5$. The CS and LCS use a time constant of 0.003 for the exponential window. For our method, we use $k = 100$ and $W = 2$.

Figure 11 shows the PSNR of each frame corrected by different methods. It can be clearly seen that the proposed method considerably reduces the FPN. Furthermore, it achieves minor residual error and faster convergence when compared with the other two NUC methods.

The resulting video sequence was analyzed and some sample frames of the uncorrected and corrected sequences are shown in Video 1. There are significant improvements in the amount of discernable scene details and visually the noise pattern is greatly reduced by all three methods. But the CS and LCS method introduce disturbing ghosting, which can be clearly seen by inspecting their error images (Fig. 12). However, we can hardly see any ghosting artifact in the MSCS's output and the level of the residual nonuniformity is rather low. These results are ratified by the PSNR of each image displayed.

6.2 *Applications to Real Infrared Data*

In this section, the CS, LCS, and MSCS were applied to a noisy sequence collected by using a 320 × 256 HgCdTe IRFPA camera, operating in the 8 to 14 μm range. The video contains 1500 frames captured at a rate of 25 fps. The CS and LCS use a time constant of 0.006, and the MSCS method uses $k = 50$ and $W = 2$. Video 2 and Fig. 13 show examples of the raw frame (frames 150 and 965) with the corresponding corrected versions using different methods. Before correction, most of the image details are obscured by the nonuniformity. The stripping patterns in the raw frame are most likely due to the readout architecture of the IRFPA. Notice, also, that the image appears darker in some regions as a result of the nonuniformity. Based on the spatial frequency characteristics of the pattern noise, MSCS uses $\sigma_{g\,\max} = 12$ and LCS uses a four-level LP for spectrum shaping.

From Video 2, it can be noted that the CS and LCS methods do not result in a noticeable change compared with the original image. The spatial noise has been partially removed but the underlying scene details cannot be readily seen. However, most of striping patterns are effectively removed by using the MSCS method with only some LSF nonuniformity remaining, which has a very weak influence on visual effect. The dark electric cords can be clearly seen in the corrected image of the MSCS, but blend in with the nonuniformity in the results of GCS and LCS.

For Fig. 13, the three methods have already been able to significantly reduce the FPN. However, the CS algorithm introduces disturbing ghosting artifacts, which suggests that the correction is indeed scene dependent. LCS produces a bit less ghosting artifacts than the CS method, but the presence of strange shadows (ghosting) around the lawn is also perceptible. The MSCS method effectively avoids the ghosting

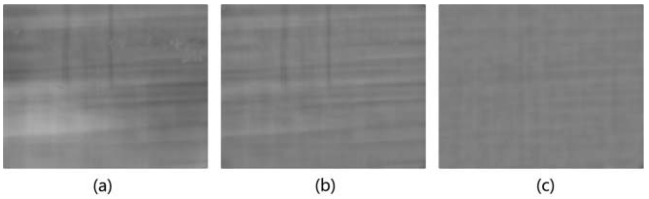
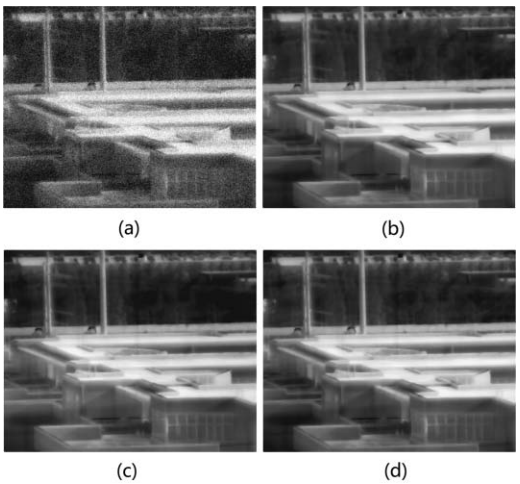
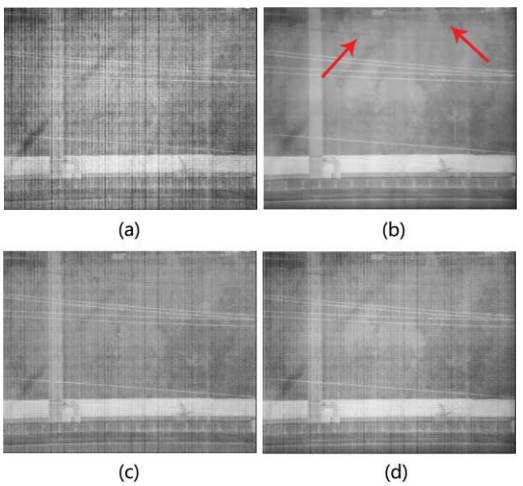


Fig. 12 Error images for (a) CS; (b) LCS; (c) MSCS. All images are scaled to the same display range.



Video 1 Simulated nonuniformity image results (frame 961). (a) Image with simulated gain and offset nonuniformity (PSNR = 23.9 dB). (b) Corrected with MSCS (PSNR = 41.9 dB). (c) Corrected with CS (PSNR = 28.2 dB). (d) Corrected with LCS (PSNR = 36.8 dB). (QuickTime, 11.46 MB).
[URL: http://dx.doi.org/10.1117/1.3610978.1]



Video 2 NUC performance comparison of the frame 150 of the real infrared sequence. (a) Unprocessed; (b) corrected with MSCS; (c) corrected with CS; (d) corrected with LCS. (QuickTime, 14.79 MB).
[URL: http://dx.doi.org/10.1117/1.3610978.2]

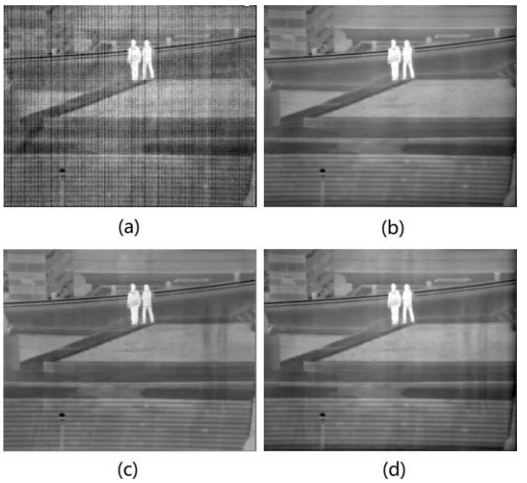


Fig. 13 NUC performance comparison of the frame 965 of the real infrared sequence. (a) Unprocessed; (b) corrected with MSCS; (c) corrected with CS; (d) corrected with LCS.

artifacts and produces, to the naked eye, a much better NUC than the other two algorithms. Through comparison, it is not difficult to find that our proposed method is superior in fast noise reduction and avoiding producing unwanted artifacts; as a result, the frame corrected by our proposed method has a better visual effect.

7 Conclusions and Discussions

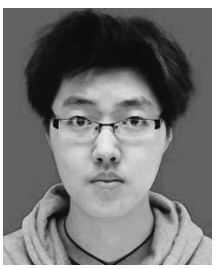
In this paper, we have presented a new statistical SBNUC method for IRFPA based on the MSCS assumption. The MSCS assumption takes regarding the relationship between the temporal signal distribution over space and the number of frame samples. It has been found that the scale that meets the constant statistics assumptions is in proportion to the number of frames approximately averaged. So, the MSCS NUC gradually expands its effective spatial scale over time according to the spatial scale that meets the CS constraint. Besides, an exponential window and a tolerance interval for the acquired data are introduced to capture the drift in nonuniformity and eliminate the ghosting artifacts. After several tests using both synthetic and real infrared data, our proposal not only performs an efficient nonuniformity correction of the sequences, but it also produces compensated images with better quality than the GCS method and the LCS method. More importantly, our method is quite simple utilizing low computational resources and very little memory, which makes it more competitive in real-time processing.

Acknowledgments

This work was supported by the Research and Innovation Plan for Graduate Students of Jiangsu Higher Education Institutions, China (Grant No.CXZZ11_0237).

References

1. D. A. Scribner, M. Kruer, and J. Killiany, "Infrared focal plane array technology," *Proc. IEEE* **79**, 66–85 (1991).
2. D. Scribner, K. Sarkady, J. Caulfield, M. Kruer, G. Katz, and C. Gridley. "Non-uniformity correction for staring focal plane arrays using scene-based techniques," *Proc. SPIE* **1308**, 224–233 (1990).
3. O. Riou, S. Berrebi, and P. Bremond, "Nonuniformity correction and thermal drift compensation of thermal infrared camera," *Proc. IEEE* **5405**, 294–302 (2004).
4. J. Harris and Y. Chiang, "Non-uniformity correction of infrared image sequences using the constant-statistics constraint," *IEEE Trans. Image Process.* **8**(8), 1148–1151 (1999).
5. J. Harris and Y. Chiang, "Minimizing the 'ghosting' artifact in scene-based nonuniformity correction," *Proc. SPIE* **3377**, 106–113 (1998).
6. C. Zuo, Q. Chen, G. Gu, and W. Qian, "New temporal high-pass filter nonuniformity correction based on bilateral filter," *Opt. Rev.* **18**, 192–202 (2011).
7. C. Zhang and W. Zhao, "Scene-based nonuniformity correction using local constant statistics," *J. Opt. Soc. Am. A* **25**, 1444–1453 (2008).
8. R. C. Hardie, M. M. Hayat, E. E. Armstrong, and B. Yasuda, "Scene based non-uniformity correction using video sequences and registration," *Appl. Opt.* **39**, 1241–1250 (2000).
9. B. M. Ratliff, M. M. Hayat, and R. C. Hardie, "An algebraic algorithm for nonuniformity correction in focal-plane arrays," *J. Opt. Soc. Am. A* **19**, 1737–1747 (2002).
10. C. Zuo, Q. Chen, G. Gu, and X. Sui, "Scene-based nonuniformity correction algorithm based on interframe registration," *J. Opt. Soc. Am. A* **28**, 1164–1176 (2011).
11. D. L. Perry, "Linear theory of nonuniformity correction in Platinum Silicide focal plane arrays," *Proc. SPIE* **1762**, 60 (1993).
12. P. J. Burt, "Fast filter transforms for image processing," *Comput. Graph. Image Process.* **16**, 20–51 (1981).
13. P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.* **31**, 532–540 (1983).
14. E. Vera and S. Torres, "Fast adaptive nonuniformity correction for infrared focal-plane array detectors," *EURASIP J. Appl. Signal Process.* **13**, 106–117 (2005).
15. T. Zhang and Y. Shi, "Edge-directed adaptive nonuniformity correction for staring infrared plane arrays," *Opt. Eng.* **45**(1), 016402-1 (2006).
16. A. Rossi, M. Diani, and G. Corsini, "Bilateral filter-based adaptive nonuniformity correction or infrared focal-plane array systems," *Opt. Eng.* **49**(5), 057003-1–057003-13 (2010).
17. A. Rossi, M. Diani, and G. Corsini, "Temporal statistics de-ghosting for adaptive non-uniformity correction in infrared focal plane arrays," *Electron. Lett.* **46**(5), 348–349 (2010).
18. J. Bai, Q. Chen, W. Qian, and X. Wang, "Ghosting reduction in scene-based nonuniformity correction of infrared image sequences," *Chin. Opt. Lett.* **8**, 1113–1116 (2010).
19. D. A. Scribner and K. A. Sarkady, "Test and evaluation of stability in IR staring infrared focal plane arrays after nonuniformity correction," *Proc. SPIE* **1108**, 255–264 (1989).



Chao Zuo received his BS degree from Nanjing University of Science and Technology, People's Republic of China, in 2009. He is currently pursuing his PhD degree in the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology, People's Republic of China. His research interests include signal and image processing, algorithms for infrared spectral sensors and imagers, information optics and digital holography. He is student member of IEEE and SPIE.

Qian Chen received his BS, MS, and PhD degree from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. He is currently a professor and the dean of the department of optoelectronic technology, Nanjing University of Science and Technology. He has led many research projects and authored more than 100 journal papers. His works have covered different topics, such as real-time digital image processing, optoelectronic imaging, electro-optical displaying technology, optoelectronic signal processing and transmission. He is member of SPIE and SID.

Guohua Gu received his BS, MS, and PhD degrees from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. He is currently a professor and director of the department of optoelectronic technology, Nanjing University of Science and Technology. He is interested in optoelectronic signal processing, transmission and infrared engineering. He is member of SPIE.

Xiubao Sui received his PhD degree in optical engineering from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. He is interested in the driver of infrared focal plane arrays and the theory research of infrared detectors.

Weixian Qian received his PhD degree in optical engineering from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. Now his research interests include infrared image processing and target tracking and detection.