Scene-Based Non-Uniformity Correction with Readout Noise Compensation

Martin Bürker[1]([✉]) and Hendrik P.A. Lensch[2]

[1] Daimler AG, Wilhelm-Runge-Straße 11, 89081 Ulm, Germany
martin.buerker@daimler.com
[2] Department of Computer Science, Computer Graphics,
Tübingen University, Sand 14, 72076 Tübingen, Germany

Abstract. Thermal cameras can not be calibrated as easily as RGB cameras, since their noise characteristics change over time; thus scene-based non-uniformity correction (SBNUC) has been developed. We present a method to boost the convergence of these algorithms by removing the readout noise form the image before it is processed. The readout noise can be estimated by capturing a series of pictures with varying exposure times, fitting a line for each pixel and thereby estimating the bias of the pixel. When this is subtracted from the image a noticeable portion of the noise is compensated. We compare the results of two common SBNUC algorithms with and without this compensation. The mean average error improves by several orders of magnitude, which allows faster convergence with smaller step sizes. The readout noise compensation (RNC) can be used to improve the performance of any SBNUC approach.

1   Introduction

In thermal cameras the noise is both spatially varying and the noise characteristics change over time. Hence, it is necessary to constantly update the noise estimate. Scene-based non-uniformity correction (SBNUC) estimates the noise parameters of the video stream online.

We present a method to estimate the readout noise by capturing a series of images with different exposure times. Adding readout noise compensation (RNC) to existing SBNUC techniques increases the initial image quality drastically.

The variety of SBNUC techniques can be devided in three major groups: Neural net (NN) algorithms assuming neighboring pixels to capture the same brightness, constant statistics/constant range (CR) algorithms assuming a uniform distribution of a pixel's values over time and motion constraint algorithms assuming a scene point can be found at a different location in the following frame. All of these approaches start from an unnecessarily noisy image due to the contained readout noise.

By removing this fixed pattern noise (FPN) the SBNUC algorithm can work on images with less corruption, which leads to faster convergence and better results. We compared NN and CR algorithms with and without RNC to validate the advantage of this method. Since the step size of the adjustment can be reduced it also has beneficial effects on ghosting.

2   Previous Work

The choice of a photometric calibration process for a digital cameras is depending on the wavelength the sensor is sensitive to. For conventional RGB cameras and some infrared cameras the procedure requires the user to take dark frames, where the lens is covered to let no light fall onto the sensor, and flat-fields, where every pixel is illuminated with the same amount of brightness (see Granados et al. [2] for details). For cameras sensitive to mid- and far-infrared it is not possible to take dark frames, since the material to cover the lens would have to be cooled to 0 K or it will emit noticeable radiation.

To overcome this problem special calibration devices called black body radiators have been developed. They have a spatially uniform area of material with the same temperature. The dark frame can be extrapolated by taking measurements at different temperatures.

One of the first methods for SBNUC has been developed by Scribner et al. [9]. They proposed to learn the gain and offset by minimizing the difference to the mean value of the nearest neighbors with steepest descent. This approach is inspired by the way neural nets backpropagate errors to adopt to the desired output. Harris and Chiang [3] assume that the brightness corresponding to a fixed pixel over time follows a Gaussian distribution with zero mean and unit variance. They called it constant-statistics constraint and used it to level out the mean and standard variation of the detector. Geng et al. [1] improved the algorithm by combining the Gaussian kernel with a temporal median. The median filter adds robustness to variation on the sample distribution. Hayat et al. [4] assume a uniform distribution of the signal. The incident radiation is estimated by using the mean and variance of the pixel values in a time window. The minimum and maximum values of the image in that window are used to project the values back to the original domain. This algorithm was refined by Torres, Reeves and Hayat [11] with a recursive method for updating the parameters of the constant range algorithm, which in turn was enhanced by Pezoa et al. [5] by replacing the simple moving average with an exponential moving average. Torres and Hayat [10] also developed a Kalman filter considering gain and offset of the detectors as state variables modeled by a Gauss-Markov process. The Kalman filter approach was then modified by San Martin, Torres and Pezoa [7] assuming that the gain does not change over time and only the offset has to be estimated. A more recent development are interframe registration based methods, like the one proposed by Zuo et al. [14]. They find the translation between two consecutive frames by computing the cross-power spectrum. The phase correlation is limited to pure translations, so this method it best suited for scanning applications where the scene moves parallel to the image plane, or vice versa.

3   Noise Model

There are two different kinds of noise cameras are susceptible to: temporal and spatial. While temporal noise can be easily dealt with by averaging multiple

frames, it is often impossible to do so due to motion in the scene or the limited frame rate of the camera. Spatial noise, or non-uniformity is the deviation of the response of each pixel to the same signal.

3.1 Temporal Noise

The variation of a pixels value exposed to the same signal over time is called temporal noise. It is caused by small variations in the conversion of light into electrons and holes (photon shot noise), the temporal variations of electrons and holes generated by the sensors temperature (dark current shot noise) and the noise of the electronic device itself occurring during the charge-to-voltage transfer and analog-to-digital conversion (readout noise).

As stated before, temporal noise can be compensated by averaging multiple consecutive frames. If this not practical the noise is often suppressed by low-pass filtering or more advanced methods in a companion chip in the camera. For calibration purposes, one wants to make sure such noise compensation is turned off since it adds unwanted non-linearity.

3.2 Spatial Noise

The quantum efficiency of the pixels is not uniform throughout the sensor i.e., for the same signal the digital value differs from pixel to pixel. This is called spatial noise and compensated by photometric calibration also known as non-uniformity correction.

Photo-Response Non-uniformity (PRNU). Each pixel consist of a photo-sensitive area in which the light is converted into current. The current is converted into voltage, which is amplified to make the result less susceptible to noise in the readout process. Due to small differences in size and material there are not two identical pixels and hence the resulting value is not the same for the same signal. This is called photo-response non-uniformity and simplified to a per pixel gain factor.

Dark Current Non-uniformity (DCNU). As each pixel responds different to light it also reacts different to temperature. The amount of current generated by temperature is called dark current and the different susceptibility to it is called dark current non-uniformity. It is modeled as an additive offset to the signal.

4 Photometric Calibration

4.1 Camera Model

A camera converts light into digital values. The irradiance $X^{(j)}$ at a pixel position j generates a digital value

$$Y^{(j)} = gt(a^{(j)}X^{(j)} + b^{(j)}) + N_R,\tag{1}$$
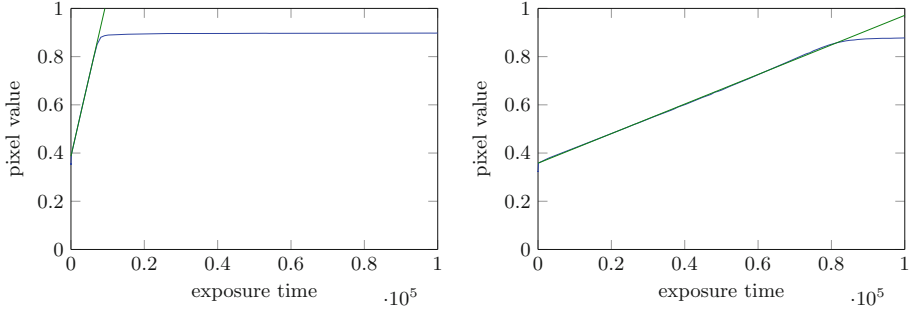
Fig. 1. Two pixels at different positions. On the left a pixel capturing a bright signal; on the right a pixel capturing a dark signal. The blue line represents the pixel value for each exposure time with 100 sampling points. The green line is the estimate of the linear regression (LMS). The values are scaled from 14 bits per pixel to $[0, 1]$, the exposure time is in μs (Color figure online).

where g is a global gain factor, t is the exposure time, $a^{(j)}$ is the per pixel gain induced by PRNU, $b^{(j)}$ is the offset induced by DCNU and N_R is the readout noise. The camera response is assumed to be linear and, unlike in [2], the quantization is omitted.

4.2   Non-uniformity Correction (NUC)

Most NUC algorithms drop the readout noise N_R and global gain g which simplifies the equation and allows to compute the irradiance $X^{(j)}$ (or some value proportional to it) by estimating the per-pixel gain $a^{(j)}$ and offset $b^{(j)}$, by

$$X^{(j)} \sim \frac{Y^{(j)}/t - b^{(j)}}{a^{(j)}}. \tag{2}$$

As we will show it is not advisable to omit the readout noise. It is neither negligible, nor hard to measure.

5   Readout Noise Compensation

In theory the readout noise is measured by taking a bias frame. That is an image captured with zero integration time which according to the camera model (Eq. 1) would consist only of readout noise. The problem is, no camera can reset and readout the sensor with no delay. This means there is always dark current present in the picture. One way is to set the integration time to the minimum value and cover the lens, as one would taking a dark frame. The captured image should contain only readout noise and very little dark current. We found that the cameras we have access to don't allow a integration time of zero and even very small values were not applied correctly.

(a) Input Y

(b) r_a (values increased by $\times 10^4$)


(c) r_b

(d) $Y - r_b$
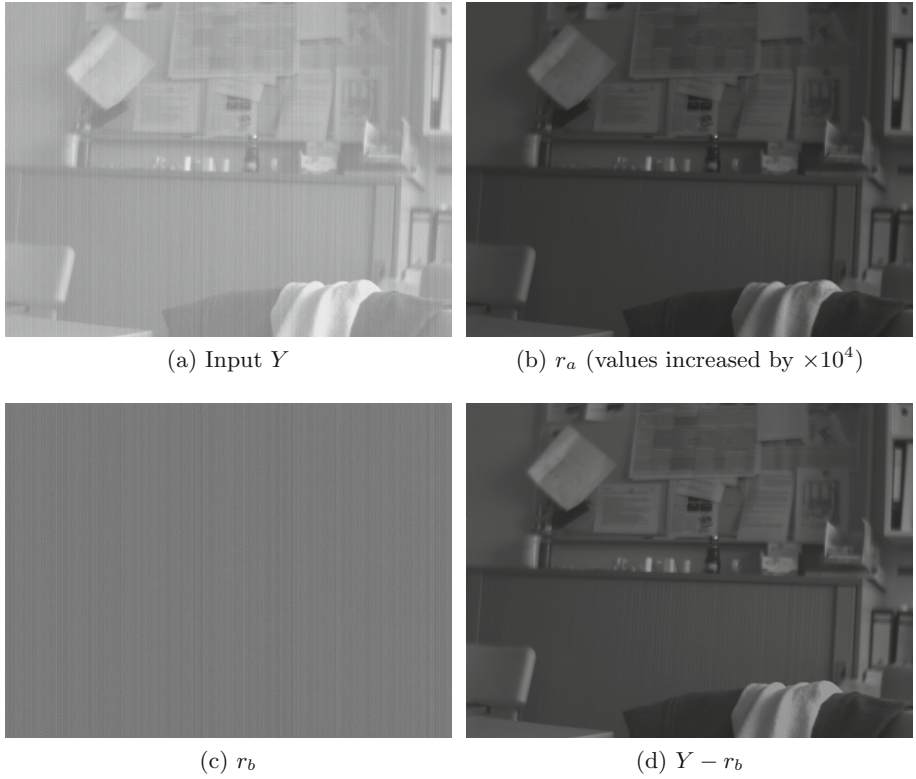
Fig. 2. The image (a) can be preprocessed by estimating the readout noise (c) and subtracting it from the input.

Another way is to estimate a pixel value with zero integration time by linear regression. This is only possible if the camera response curve is fairly regular i.e., the digital value should change with the same factor as the integration time (Fig. 1). For a single pixel j the same scene point is captured with different exposure times $t_1, t_2, ..., t_n$, so the pixel values are

$$\begin{pmatrix} Y_1^{(j)} \\ Y_2^{(j)} \\ \vdots \\ Y_n^{(j)} \end{pmatrix} = r_a^{(j)} \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} + r_b^{(j)}, \tag{3}$$

where r_a and r_b can be estimated by linear regression e.g., least mean square. To get more accurate results, the temporal noise can be reduced by averaging multiple exposures or the number of integration times can be increased. Note that it is not necessary to know the signal $X^{(j)}$ or any of the other variables in
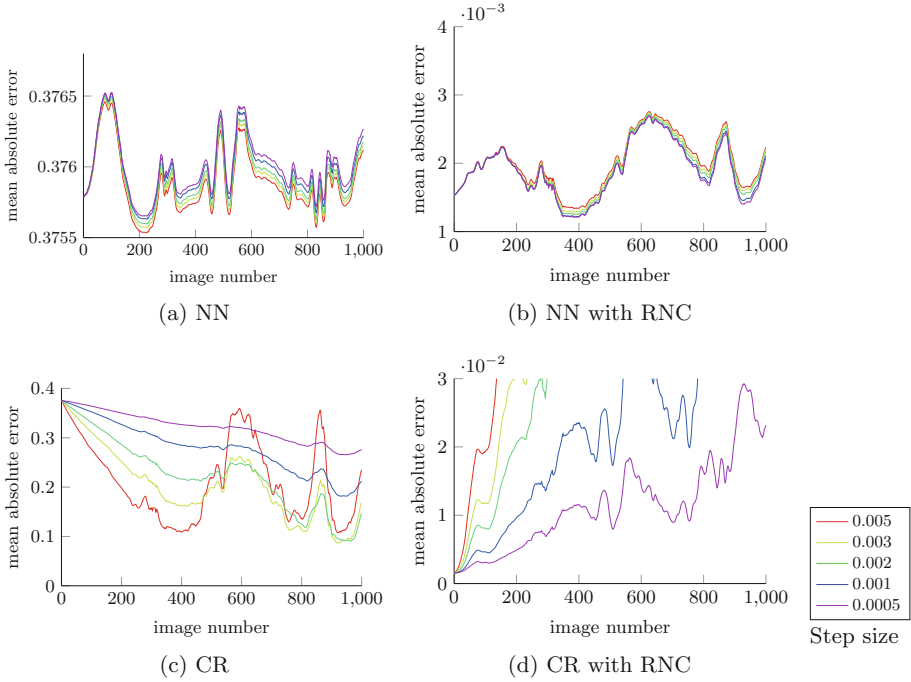
Fig. 3. The mean absolute error (MAE) to the ground truth of the Xenics Bobcat for different step sizes.

Eq. 1, since the parameters

$$
\begin{aligned}
r_a^{(j)} &= g(a^{(j)} X^{(j)} + b^{(j)}), \text{ and} \\
r_b^{(j)} &= N_R
\end{aligned}
\tag{4}
$$

already comprise them. It is essential that none of the values $Y_i^{(j)}$ is saturated.

We can now simply remove the readout noise by subtracting $r_b^{(j)}$ from the pixel value $Y^{(j)}$ (see Fig. 2).

Readout noise is classified as temporal noise, which means it should not have a different average for different pixels. This means either that there is also a spatial non-uniformity in readout noise, or that the non-uniformity we found is caused by a noise source which our camera model does not contain.

6   Evaluation

Our readout noise compensation (RNC) has been evaluated on two common SBNUC algorithms: neural nets (Scribner et al. [8]) and constant range (Hayat et al. [4]). Both algorithms were implemented on a graphics card with the following improvements:
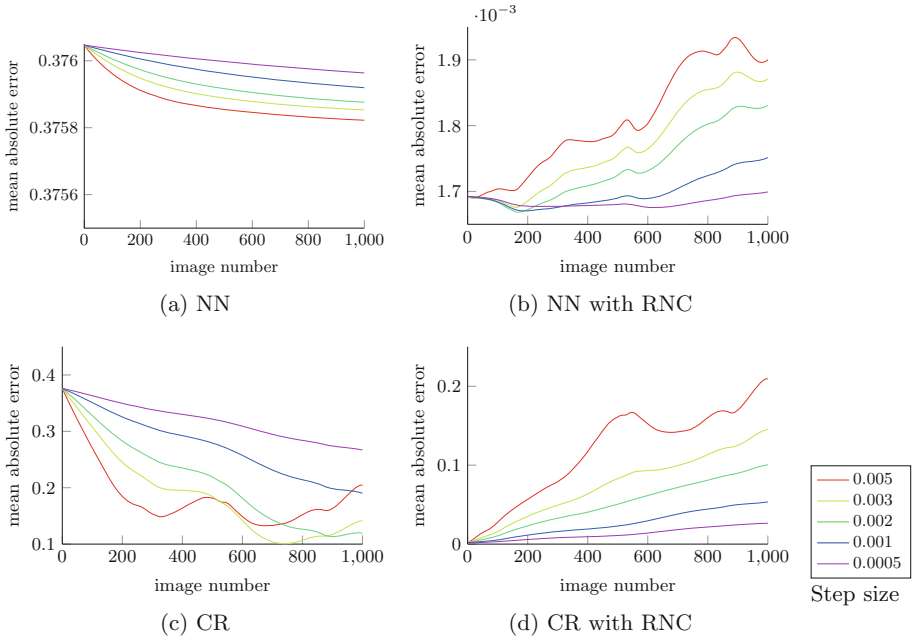
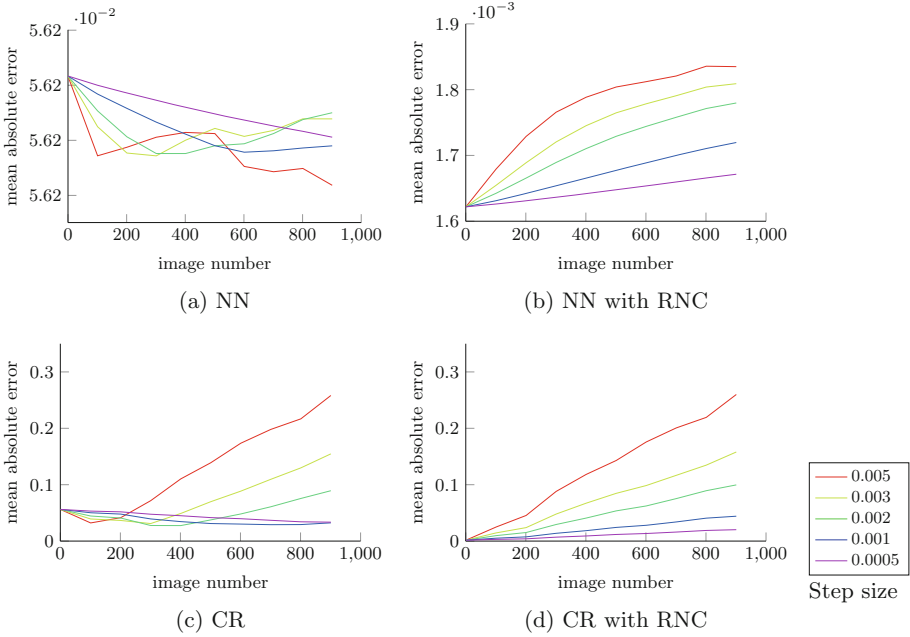**Fig. 4.** The mean absolute error to the evaluation data of the Xenics Bobcat for different step sizes.

- The neural net algorithm was implemented with the adaptive learning rate from Torres et al. [12].
- The constant range approach was implemented according to Redlich et al. [6]. The minimum and maximum of the range were set per pixel to the lowest and highest value of the frames processed.

A Xenics Bobcat-640-CL and a Raptor Photonics OWL 640 CameraLink were used for the evaluation. The sequences used were captured with the cameras mounted behind the (uncoated) windshield of a car. Due to the different frame rates, the length of the sequences differ from each other, so we took the first 1000 frames of each sequence. The images were processed in the order they were captured in to simulate an online calibration.

The results are heavily depending on the chosen step size. If it is too small, the algorithm would take longer than needed; if they are too big the algorithm would produce ghosting (overfitting). In general it is advisable to decrease the learning rate over time, which is called annealing (see Zeiler [13]), but for comparison it is sufficient to use a fixed step size.

6.1   Comparison to Ground Truth

The error was measured in mean absolute error (MAE) to a ground truth. The ground truth is the image corrected by flatfielding and dark frame subtraction.

(a) NN

(b) NN with RNC

(c) CR

(d) CR with RNC

Fig. 5. The mean absolute error to the evaluation data of the Raptor OWL for different step sizes.

For each camera several 50 flatfields and 50 dark frames at operating temperature were acquired. The mean flatfield ff and mean dark frame b are used to calculate the per-pixel gain (similar to [2])

$$a^{(j)} = \frac{ff^{(j)} - b^{(j)}}{1/m \sum_{j=1}^{m} \left(ff^{(j)} - b^{(j)} \right)}. \tag{5}$$

The ground truth is then computed according to Eq. 2. Since it is not used in any of the evaluation algorithms, the global gain is not corrected for. For a valid ground truth it is absolutely necessary to capture dark frames. This is only possible with cameras sensitive to reflected and not thermal light, which is why we used short-wave infrared cameras for the evaluation.

6.2   Comparison to Evaluation Data Set

The result of each frame is also compared to a random (but fixed) set of 30 images from other sequences. This tests the generality of the current gain and offset. The overfitting to pixels that change only slightly is penalized when choosing uncorrelated images to compute the MAE.

(a) NN

(b) NN with RNC

(c) CR

(d) CR with RNC
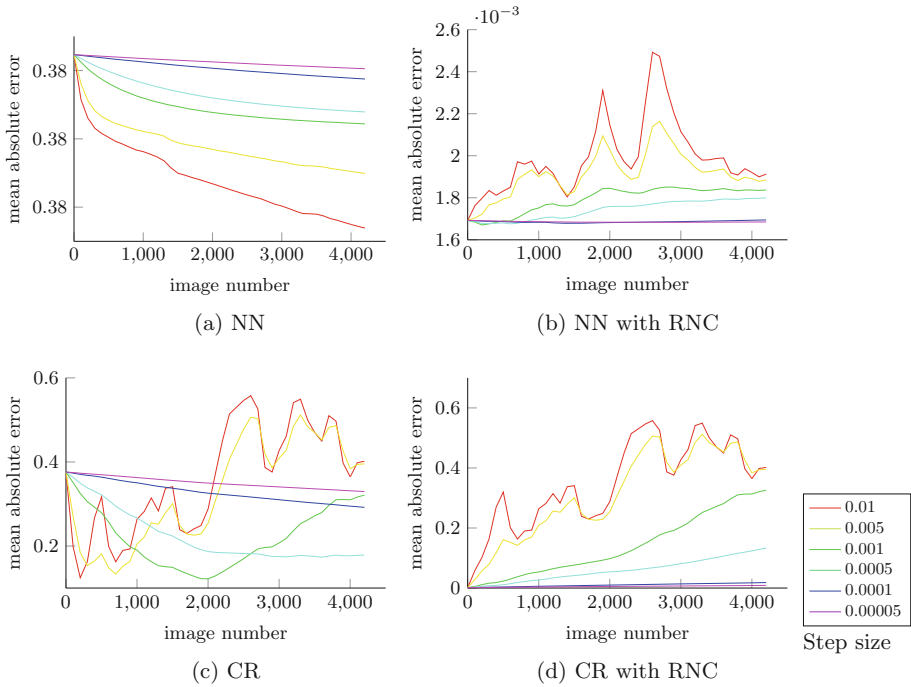
Fig. 6. The mean absolute error to the evaluation data of the Xenics Bobcat for different step sizes.

6.3   Longer Sequence

To make sure we do not stop the evaluation before the algorithms could exploit their full potential, we used a 4235 frame sequence of the Xenics Bobcat (Fig. 6). Also we used a wider range of step sizes with higher values to give the NN the chance to take bigger steps towards a smooth image and smaller values to let the algorithms with RNC slowly decrease the MAE even more.

6.4   Results

As Figs. 3, 4 and 5 show, without RNC the algorithms could not achieve an MAE even close to the error of the pictures with RNC after 1000 frames. The MAE achieved with RNC is around 0.16 %, whereas the lowest MAE achieved without RNC is 9.1 % (CR to ground truth), or 11.7 % (CR to evaluation data). As Fig. 6 shows, even after 4235 frames none of the algorithms could decrease the MAE significantly under the initial error of RNC. The lowest MAE to the evaluation data was achieved by NN with RNC with a step size of 0.002 after 164 frames.

(a) NN, step size 0.1

(b) NN with RNC, step size 0.001

(c) CR, step size 0.001

(d) CR with RNC, step size 10^{-6}

Fig. 7. Frame 1000 of the sequence. The results with RNC (b) and (d) have both an MAE of < 0.002. The NN algorithm (a) shows ghosting and the CR approach (c) still has visible FPN. For displaying purposes, the brightness of (b) and (d) has been increased.

7   Conclusion

We showed that the proposed RNC gives the actual SBNUC algorithm a way better starting point. Figure 7 shows that the remaining non-uniformity after 1000 frames is - with both tested algorithms - hard to see and the MAE is smaller than 0.2%. Since the RNC is independent from the SBNUC it can be also used with other algorithms than NN and CR. Even without SBNUC the RNC reduces the FPN considerably. Although, for long sequences we recommend to use SBNUC with a low step size, which can smooth out changes in dark current noise that is common in thermal cameras. In theory, once the parameters are known they can be reused unlimited, but this has not been researched and the parameters can be updated every time the camera is not moving.

References

1. Geng, L., Chen, Q., Qian, W., Zhang, Y.: Scene-based nonuniformity correction algorithm based on temporal median filter. J. Opt. Soc. Korea **17**(3), 255–261 (2013)
2. Granados, M., Ajdin, B., Wand, M., Theobalt, C., Seidel, H.P., Lensch, H.P.A.: Optimal HDR reconstruction with linear digital cameras. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 215–222. IEEE, June 2010
3. Harris, J.G., Chiang, Y.M.: Nonuniformity correction using the constant-statistics constraint: analog and digital implementations. In: Proceedings of SPIE, vol. 3061, pp. 895–905 (1997)
4. Hayat, M.M., Torres, S.N., Armstrong, E., Cain, S.C., Yasuda, B.: Statistical algorithm for nonuniformity correction in focal-plane arrays. Appl. Opt. **38**, 772–780 (1999)
5. Pezoa, J.E., Torres, S.N., Córdova, J.P., Reeves, R.A.: An enhancement to the constant range method for nonuniformity correction of infrared image sequences. In: Sanfeliu, A., Martínez Trinidad, J.F., Carrasco Ochoa, J.A. (eds.) CIARP 2004. LNCS, vol. 3287, pp. 525–532. Springer, Heidelberg (2004)
6. Redlich, R., Figueroa, M., Torres, S.N., Pezoa, J.E.: Embedded nonuniformity correction in infrared focal plane arrays using the Constant Range algorithm. Infrared Phys. Technol. **69**, 164–173 (2015)
7. San Martin, C., Torres, S., Pezoa, J.E.: Statistical recursive filtering for offset nonuniformity estimation in infrared focal-plane-array sensors. Infrared Phys. Technol. **51**(6), 564–571 (2008)
8. Scribner, D., Sarkady, K., Kruer, M., Caulfield, J., Hunt, J., Colbert, M., Descour, M.: Adaptive retina-like preprocessing for imaging detector arrays. In: IEEE International Conference on Neural Networks, pp. 1955–1960 (1993)
9. Scribner, D.A., Sarkady, K.A., Caulfield, J.T., Kruer, M.R., Katz, G., Gridley, C.J., Herman, C.: Nonuniformity correction for staring IR focal plane arrays using scene-based techniques. In: Proceedings of SPIE, Aplications of Artificial Neural Networks, vol. 1308, pp. 224–233 (1990)
10. Torres, S.N., Hayat, M.M.: Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays. J. Opt. Soc. Am. A Opt. Image Sci. Vis. **20**(3), 470–480 (2003)
11. Torres, S.N., Reeves, R.A.: Scene-based nonuniformity correction method using constant-range: performance and analysis. In: Proceedings of 6th World Multiconference on Systemics, Cybernetics and Informatics, vol. 9, pp. 224–229 (2002)
12. Torres, S.N., Vera, E.M., Reeves, R.A., Sobarzo, S.K.: Adaptive scene-based nonuniformity correction method for infrared-focal plane arrays. In: Holst, G.C. (ed.) Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5076, pp. 130–139, August 2003
13. Zeiler, M.D.: ADADELTA: An Adaptive Learning Rate Method. arXiv:1212.5701 [cs.LG] (2012)
14. Zuo, C., Chen, Q., Gu, G., Sui, X.: Scene-based nonuniformity correction algorithm based on interframe registration. J. Opt. Soc. Am. A Opt. Image Sci. Vis. **28**(6), 1164–1176 (2011)