Practical Issues with 3D-Noise Measurements and Application to Modern Infrared Sensors

Patrick O'Shea, Stephen Sousk

U. S Army RDECOM CERDEC
Night Vision and Electronic Sensors Directorate (NVESD)
Fort Belvoir, Virginia 22060-5806 USA

ABSTRACT

The two most important characteristics of every infrared imaging system are its resolution and its sensitivity. The resolution is limited by the system's Modulation Transfer Function (MTF), which is typically measurable. System sensitivity is limited by noise, which for infrared systems is usually thought of as a Noise Equivalent Temperature Difference (NETD). However, complete characterization of system noise in modern systems requires the 3D-Noise methodology (developed at NVESD), which separates the system noise into 7 orthogonal components including both temporal-varying and fixed-pattern noises. This separation of noise components is particularly relevant and important in characterizing Focal Plane Arrays (FPA), where fixed-pattern noise can dominate. Since fixed-pattern noise cannot be integrated out by post-processing or by the eye, it is more damaging to range performance than temporally-varying noise. While the 3D-Noise methodology is straightforward, there are several important practical considerations that must be accounted for in accurately measuring 3D Noise in the laboratory. This paper describes these practical considerations, the measurement procedures used in the Advanced Sensor Evaluation Facility (ASEF) at NVESD, and their application to characterizing modern and future infrared imaging systems.

Keywords: 3D Noise, NETD, FLIR, infrared imaging system, infrared system measurements

1. INTRODUCTION

The characterization of forward-looking infrared (FLIR) systems is important for several purposes, including design verification and acceptance testing, calibrating model input parameters for performance predictions, and testing for unit failure. Unlike a "typical operation" or otherwise modeled sensor parameter, which tells you how a sensor *type* typically *should* behave, a laboratory measurement tells you how a *specific* sensor system *does* behave. The two most important and fundamental characteristics of any FLIR are resolution and sensitivity. Resolution is typically characterized by the Modulation Transfer Function (MTF), the frequency representation of the system blur. This is a common and mature test for all kinds of imaging systems. For a FLIR, the only special considerations are to have a target (a sharp line or edge) that emits in the waveband of that sensor. Additionally, many modern focal plane arrays (FPA) are undersampled, and so require a measurement technique that effectively re-samples the image (scanning slit or tilted edge). Sensitivity is typically characterized as a Noise-Equivalent Temperature Difference (NETD), with units of K or mK. However, to fully characterize the noise of modern infrared sensors, especially FPA-based sensors, requires the 3D-Noise methodology[1]. The 3D-Noise method takes a sequence of frames as an image "cube," and then extracts a series of 7 orthogonal noise components, some static and some temporally varying. This is important in sensor modeling, where the directionality of the noise, and whether the noise is static or dynamic, can have a profound impact upon predicted performance. This has the strongest impact in uncooled FPA systems, which often have fixed-pattern noise that equals or exceeds the dynamic noise. Since dynamic noise is integrated by the eye, and fixed-pattern noise is not, the fixed pattern noise has a more deleterious effect on performance. More subjectively, the fixed pattern noise is more "distracting," as if trying to view an image through a dirty filter.

There are some common misunderstandings about 3D Noise, and some subtle and not-so-subtle difficulties in making the measurement and doing the analysis correctly. The purpose of this paper is to clarify some misconceptions, and address some particular practical issues in measuring and analyzing 3D Noise. These observations and solutions are a result of the ongoing measurements and measurements research of the Advanced Sensor Evaluation Facility (ASEF),

Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVI, edited by Gerald C. Holst, Proceedings of SPIE Vol. 5784
(SPIE, Bellingham, WA, 2005) • 0277-786X/05/$15 • doi: 10.1117/12.604588

the primary FLIR characterization laboratory, at the Night Vision and Electronic Sensors Directorate (NVESD), where the 3D-Noise methodology was first developed.

2. 3D-NOISE METHODOLOGY

The 3D-Noise measurement of a sensor is made by acquiring a series of digital or analog frames to create an image "cube", with dimensions of H (horizontal width, in pixels) by V (vertical image height, in pixels) by T (number of frames, or "time"). The image cube is captured usually, but not always, at the sensor frame rate. However, the measurement should be made such that the time between frames is kept constant. This volume of pixels is not always all used in computing the 3D-Noise components. For various considerations, to be detailed later, a sub-volume may be selected for the actual data reduction, with smaller values of H and V, but generally keeping the same number of frames T.

The 3D-Noise analysis is made by using a series of directional averaging operators, which when used in combination, will extract the values of each noise component from the cube, and remove that component from the original noise cube, until all the noise components have been extracted. The basic method is straightforward, but often misunderstood. There are also several shortcuts that can be made in the analysis to speed up computation times, especially for large noise cubes.

The 3D-Noise methodology is built upon some specific, but not always understood, assumptions. First, the noise components themselves have a zero mean, with Gaussian distribution. The overall data cube will almost certainly not have a zero mean, but this mean is only an overall average value *S*, which can be obtained by averaging every pixel in the cube to a single number. Additionally, the noise components are independent and uncorrelated. Given these assumptions, we can write the total measured noise cube as

$$U(t,v,h) = S + N_t(t) + N_v(v) + N_h(h) + N_{tv}(t,v) + N_{th}(t,h) + N_{vh}(v,h) + N_{tvh}(t,v,h) \qquad (1)$$

Each noise component is described in Table 1. Roughly, they can be split into three groups: the temporally-correlated spatial noise components, N_{vh}, N_v, and N_h; the spatially correlated temporal noise components, N_t, N_{th}, and N_{tv}; and the random spatio-temporal noise N_{tvh}, which is most equivalent to the traditional NETD. For FLIR modeling in NVTherm[2], the noise inputs are limited to the standard deviations of the random spatio-temporal noise and the spatial noises.

Noise	Description	Potential Source
σ_{tvh}	Random spatial-temporal noise	Detector temporal noise
σ_{tv}	Temporal row noise	Line processing, 1/*f*, readout
σ_{th}	Temporal column noise	Scanning
σ_{vh}	Random spatial noise	Pixel processing, detector nonuniformity
σ_v	Fixed row noise	Detector nonuniformity, 1/*f*
σ_h	Fixed column noise	Detector nonuniformity, scanning
σ_t	Frame-to-frame noise	Frame processing
σ_{Total}	Total Noise	RSS of All Noises

Table 1. The seven 3D-Noise components, descriptions, and possible phenomena giving rise to each component.

The noise components are extracted from the noise cube as seen in Fig. 1. The cube is operated on by the directional averaging operators D_h, D_v, and D_t, which reduce the dimensionality of the result by 1; a cube is averaged down to a plane, a plane is averaged down to a line (stream), and a line is averaged down to a single number (the global average, *S*, no matter the order of the operations). The D_A operator removes all noise components related to direction *A*. An intuitive example is capturing 128 frames and averaging all the frames to a "*VH*-plane", to smooth out transient noise.

This is exactly the D_t operator, which removes all noise terms except N_{vh}, N_v, and N_h. The operation [1-D_A] is performed by averaging the cube/frame/line along direction *A* and then subtracting the resultant frame/line/value along direction *A*

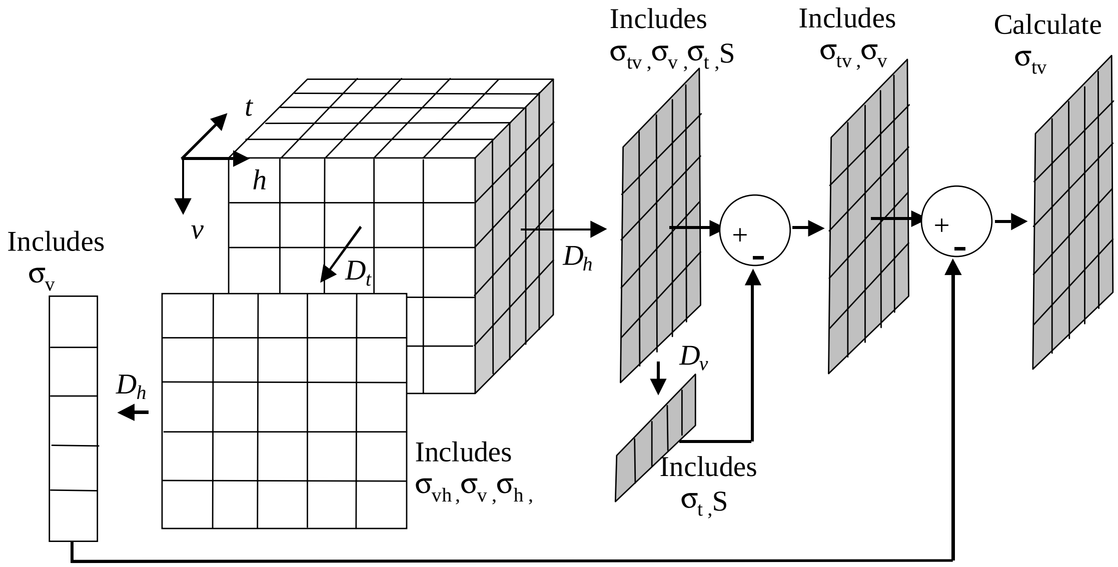


Figure 1. Sample extraction of a single noise component using successive application of the directional averaging operator.

from every frame/line/value in the original cube/frame/line. (Note that [1-*D*] does not reduce the dimensionality of the object operated on.) Extending the example, subtracting the time-averaged *VH*-plane from every frame in the original cube leaves a cube containing only the temporally varying components, N_t, N_{th}, N_{tv}, and N_{tvh}. Successive operations of *D* and [1-*D*] are performed until only one component remains. The standard deviation of that resultant object is standard deviation of the remaining noise component, such as σ_{th}, column bounce noise.

It is important to note, and often missed, that the successive averaging operations need to be performed on the successive results of the previous operations. It is computationally straightforward to average the original noise cube along the *T*, *V*, and *H* directions to get the *VH*-, *TH*- and *TV*- planes, and then subtract each result from the original to eliminate all components besides N_{tvh}. However, it is incorrect. The averaging and subtraction of the first [1-*D*] operator must be performed first, and then the second [1-*D*] operator performed on the resultant adjusted noise cube, and again with the third. The order of operations does not matter; so long as each result is passed to the next operation.

First, we note that for the singly-subscripted noises, the calculation is quite simple. Consider N_t, though the result applies to all three components:

$$\begin{aligned}
N_t &= [1 - D_t] D_v D_h U(t,v,h) \\
N_t &= D_v D_h U(t,v,h) - D_t D_v D_h U(t,v,h) \\
N_t &= D_v D_h U(t,v,h) - S \\
N_t &= U(t)
\end{aligned} \tag{1}$$

We drop the global mean value *S*, since it will not contribute to the resultant standard deviation. After averaging down in the horizontal and vertical directions, the resultant line (or "stream") *U*(*t*) is a 1-D array of values that are added to the overall values of each *t*-th frame. This noise is most intuitive in the direction *T*: it is flicker, an overall variation in brightness and darkness of the scene from frame to frame. While the noise is a 3-D component, the fact that every value in each frame is the same allows the same amount of information to be reduced to a 1-D stream. Subtraction of streams from cubes simply requires subtracting a given frame's value from every pixel of that frame, along the direction of the stream for each value of that stream.

However, for N_{tvh}, the calculation must be more careful. This can be seen by multiplying out the D operators:

$$
\begin{aligned}
N_{tvh} &= [1-D_t][1-D_v][1-D_h]U \\
N_{tvh} &= U - D_tU - D_vU - D_hU + D_tD_vU + D_tD_hU + D_vD_hU - D_tD_vD_hU \\
N_{tvh} &= \{U - D_tU - D_vU - D_hU\} + \{D_tD_vU + D_tD_hU + D_vD_hU\} - S \\
\{U - D_tU - D_vU - D_hU\} &= N_{tvh} - \{N_h + N_v + N_t\}
\end{aligned}
\tag{2}
$$

where U is implicitly $U(t,v,h)$. The global mean value S can be dropped, since it will not contribute to the resultant standard deviation. The left-hand side of Eq. 2 is the resultant cube made from separate but not successive [1-D] operations. The right-hand side of Eq. 2 shows that the result is not N_{tvh}, but oversubtracts each doubly-averaged line component. However, it may be computationally efficient to perform all the averaging, first to planes and then to lines, and perform all the subtractions, so long as the oversubtracted terms are added back in to correct for them. Note that this oversubtraction does not present a problem for any of the other noise components. Consider N_{vh}, though the result can be generalized by simply permuting the indices:

$$
\begin{aligned}
N_{vh} &= [1-D_v][1-D_h]D_tU \\
N_{vh} &= [1-D_v][1-D_h]U_{vh} \\
N_{vh} &= U_{vh} - D_vU_{vh} - D_hU_{vh} + D_vD_hU_{vh} \\
N_{vh} &= U_{vh} - D_vU_{vh} - D_hU_{vh} = U_{vh} - N_h - N_v
\end{aligned}
\tag{3}
$$

where U_{vh} is the time-averaged VH-plane, and recognizing that $D_vD_hU_{vh} = D_vD_hD_tU = S$, which can again be dropped. What remains is identical averaging the VH-plane vertically to get N_h, averaging the VH-plane horizontally to get N_v, and subtracting each result from the original VH-plane. The only extra term that appears from doing the calculation "right," with successive averaging of the previously corrected frame, is the global average S, which doesn't alter the variance of the noise (which is likely why many people fall into the trap of doing it "wrong" for N_{tvh}).

A figure of merit sometimes used is the "Spatial NETD" or "Spatial Noise." This often is just the standard deviation of the time averaged frame (U_{vh}). It is equivalent to

$$
\sigma_{spatial} = \sqrt{\sigma_{vh}^2 + \sigma_v^2 + \sigma_h^2}
\tag{4}
$$

Another metric sometimes quoted is "Temporal Noise" or "Temporal NETD." This is often found by finding the standard deviation along T, for each pixel in the VH dimension, and then averaging the V*H results:

$$
\begin{aligned}
\sigma_{temporal}(v,h) &= \sqrt{\frac{1}{T}\left(\sum_t U^2_{t,v,h} - \left(\sum_t U_{t,v,h}\right)^2\right)} \\
\sigma_{temporal} &= \frac{1}{VH}\sum_{v,h}\sigma_{temporal}(v,h)
\end{aligned}
\tag{5}
$$

This is really a per-pixel measure of variance, and can be thought of as an average individual processed detector noise. It is not really a measure of overall scene temporal noise, which would instead be something like

$$
\sigma_{temporal\_scene} = \sqrt{\sigma_{tvh}^2 + \sigma_{tv}^2 + \sigma_{th}^2 + \sigma_t^2} = \sqrt{\sigma_{total}^2 - \sigma_{spatial}^2}
\tag{6}
$$

where σ_{total} is the standard deviation of the raw noise cube, which would be:

$$\sigma_{total} = \sqrt{\sigma_{tvh}^2 + \sigma_{tv}^2 + \sigma_{th}^2 + \sigma_t^2 + \sigma_{vh}^2 + \sigma_v^2 + \sigma_h^2} \tag{7}$$

There is one last assumption to keep in mind. Since the noise is assumed to be of zero mean, the expectation is that the averaging operator will sum up noises along that direction to zero. In practice, of course, each averaged component does not necessarily sum to zero, but is expected to be approximately as small as the standard error, $\sigma/\sqrt{N}$. This causes an oversubtraction in the use of the D operators, making σ_{tv}, σ_{th}, and σ_{vh} slightly larger at the expense of σ_{tvh}, but then also makes σ_t, σ_v, and σ_h slightly larger at the expense of σ_{tv}, σ_{th}, and σ_{vh}. To minimize this error, N should be kept large, 100 or more. There is a tradeoff involved, since large values of N often require large amounts of data storage, especially for digital data of large-format FPAs. One can alter the 3D-Noise algorithm to attempt to account for this discrepancy, but such a treatment is beyond the scope of this paper.

3. 3D-NOISE MEASUREMENTS

There are several practical considerations in the actual laboratory measurement of 3D Noise. Some of the issues are intrinsic to the measurement, such as bit depth, frame rate and data storage. Others are tied into the methodology, such as the size of the image cubes captured and analyzed. Some of the issues are particular to analog capture, when digital image capture is not readily available. Some are implementation specific, such as edge enhancement and its effect on calibration. While not every conceivable issue can be covered in any document, the most common ones will be covered, as well as a few less common, but subtle concerns.

It is always preferable to acquire digital data, if readily available. This traditionally takes twice the space of 8-bit analog data, and consequently affects bandwidth required on the capture hardware and software, as well as nominally doubling the size of the captured images, before any compression (obviously, any such compression must be absolutely lossless). Depending on the hardware and software performing the data acquisition of image frames, and the array size, there may be some limit on the number of frames that can be captured, especially for large-format arrays. As described in the methodology section, there is an associated error in finite data sets that goes as $1/\sqrt{N}$. A good rule of thumb is to take at least 100 frames of data, keeping that error to, at most, about 10% for σ_t. This error will always lower the value of σ_{tvh}, and on average, the fewer the frames, the lower σ_{tvh} is, though the total RSS noise will not change. Due to the unfortunate preponderance of "monomania[†]," the most quoted number of the 3D Noise components is σ_{tvh}, and is commonly misrepresented as "the NETD." As human nature is beyond the control of the measurement, it behooves the honest measurements laboratory to minimize this erroneous reduction in σ_{tvh}, and to have at least 100 frames of data. Additionally, since the result of the 3D Noise measurement and analysis is a series of standard deviations that are supposed to fully represent the sensor, it is good practice that the noise cubes be at least 100x100x100 in size, just to sample well enough to effectively minimize sampling error in the statistics.

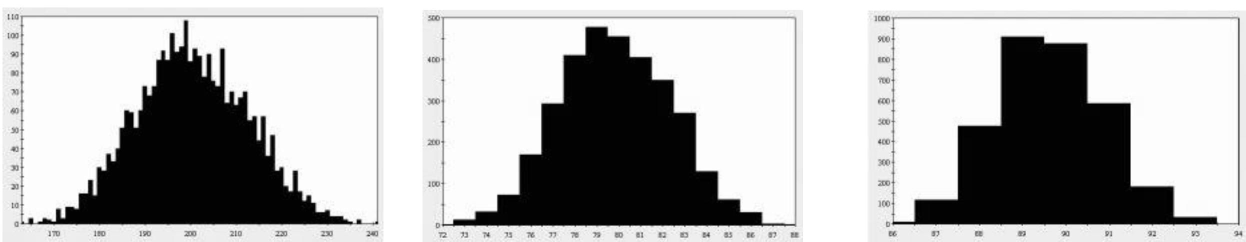
If digital data is being acquired, one can almost always rely upon the direct response of output voltage to calibrated temperature. This allows the acquisition of "full frame" 3D Noise data, by flooding the field of view (FOV) of the sensor with a "spatially uniform" flat-plate blackbody (hereafter called just "blackbody") source. There is no temperature reference in the scene, but the temperature of the blackbody can be lowered above and below ambient to map out a Signal Intensity Transfer Function (SITF), and calibrate the system, mapping change in input temperature to change in output signal (such as greyshade/mK). Note that "spatially uniform" is relative, and is more stringent for highly uniform sensor arrays, but most commercially available blackbodies are sufficient for even the most spatially uniform sensor arrays. One must take care, of course, that the blackbody has time to relax and equilibrate spatially, especially after large temperature jumps, before beginning any measurement that depends on image uniformity (which applies to *all* infrared characterization measurements).

Once the calibration is acquired via the SITF, a digital noise cube can be acquired, of the full horizontal and vertical extent of the sensor, with many frames of data. For 128 frames of 16-bit digital data (the useful data may of course be of a lower bit depth, say 10- or 12- bit) a typical 640x480 FPA, this takes 128 frames * 2 bytes * 640 * 480 = 75 MB of

[†] The tendency to oversimplify measurement results by attempting to reduce them to a single number.

data, which must be stored into free memory or streamed directly to a digital storage medium. The 3D-Noise algorithm is used to extract the seven noise components from the image cube, the standard deviations of the components calculated, and then the standard deviations converted to temperature via the SITF-derived calibration.

Depending on the FPA gain and electronics, the values for each noise component (which are representable as lines for N_v, N_h, and N_t; planes for N_{tv}, N_{th}, or N_{vh}; or cubes for N_{tvh}) will be "binned" either coarsely or finely, as seen in Fig. 2. If the FPA gain is quite low, the system will have a good dynamic range, but this can make the "binning" quite coarse, with the number of different distinct digital values quite small. The noise is assumed to be Gaussian distributed, but very coarsely binned histograms of the noise components make poor Gaussians, and the meaning, accuracy, and value of the standard deviation can be affected. It is uncommon for sensors to have easy access to FPA (pre-amplifier) gain; if the binning in digital data is coarse (say, the bulk of all counts in only a few bins), not much can be done to alter the binning. Two approaches to consider are taking multiple 3D Noise cubes, even at the same conditions; and to shift the temperature of the scene, in essence shifting the "phase" of the (hopefully) Gaussian-distributed noise with respect to the bins, and taking multiple measurements. This difficulty and solution to it is identical in analog measurements, if the sensor is already in maximum gain. If not, the analog gain can be increased to spread the noise over more grey levels, which will result in one of two things: removing the gross coarseness of the binning, solving the problem, or merely stretching the few digital bins into a comb of separated analog bins, which, while increasing contrast to the eye, add no new statistical information; and one goes back to the digitally coarse problem and solution.



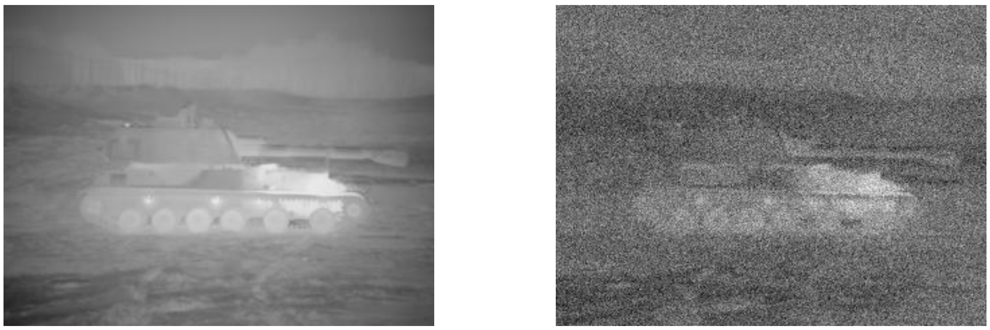
(a) (b) (c)

Figure 2. Target area histograms with (a) high gain and fine "binning" through (c) low gain and coarse "binning."

One critical assumption of the 3D Noise algorithm is that all the noises have a nominally zero mean. This is a good assumption for the temporally varying components over the course of the measurement. If it were not true, then the total brightness of the scene would be either brighter or darker from the beginning to the end of the measurement, which is typically only a few seconds long. If that trend were to continue, then the scene would go either completely black or completely white in a short period of time. If the trend were to cycle, say with a period of 20 seconds or more, then a single measurement would not pick up that this noise has a zero mean. If the cycle is less than the measurement time, but many frame times, then this is a low frequency temporally varying noise. The usual interpretation of the 3D Noise components also assumes the noise has a flat spectrum. A full 3D Noise Power Spectral Density measurement would be required to fully characterize such a low-frequency noise, which is beyond the scope of this paper. However, such a temporally varying noise component is quite rare, and as such, the zero-mean assumption is quite good, and so any standard deviation in the temporally varying components is almost certainly a measure of the standard deviation of the noise. Additionally, in those cases where there is a temporally varying component with a clear frequency, it is intuitively obvious to the observer, and the frequency will often give some indication of the underlying cause, usually a design flaw or component failure.

The same assumptions can not be made about the spatial noise components. For example, the incoming light focused by the sensor will generally have a curved wavefront onto a flat focal plane, causing optical rolloff of signal, which the Non-Uniformity Correction algorithm of the sensor may or may not correct. If an internal flag is used at the uniform scene for the NUC, then the optical rolloff will not be accounted for. What this can mean, especially for fast optical systems, is there can be a significant spatial variation in the scene, even if the FPA itself had no noise at all!

What's going on here is a semantic difference between "spatial noise" and nonuniformity. Nonuniformity can be viewed as just low-frequency spatial noise. The distinction is important, however, when it comes to sensor performance. If a sensor with nearly-zero detector and amplifier noise has a clear nonuniformity, the standard deviation of the total spatial noise could be quite large, such as the image in Fig 3a. But, if an "equivalent" sensor existed with no nonuniformity, but had zero-mean, Gaussian-distributed spatial noise totaling to the same large standard deviation, the sensor performance would be significantly worse, as seen in Fig. 3b. This is true because the target does not contain frequencies as low as the nonuniformity, and noise outside the spatial frequency of the target has minimal impact on performance[3]. This, in fact, is what happens if the noise is blithely entered into performance models, without the associated noise spectrum, and the sensor is unfairly penalized. For this reason, it is always important to remove the low-frequency trends in the temporally averaged spatial noise frame (the "*VH*-plane"). For the reasons described in the previous paragraph, it is not proper to do this in the temporally-varying dimensions of the noise cube; any trend is assumed to be real and detrimental to performance.



(a) (b)

Figure 3. (a) A scene with an overall nonuniformity of a large standard deviation, but with minimal impact on task performance. (b) The same scene with white σ_{vh} noise of the same standard deviation, but with a much more deleterious effect on task performance.

To remove only low-frequency variations in spatial noise, a low-order two-dimensional least-squares polynomial fit is used to find the trend and remove it from the time-averaged frame. Usually a second-order fit is sufficient and appropriate, but if the cause of the nonuniformity is known to require a higher-order polynomial to remove, then one should be used. It may be computationally efficient and adequate to perform, line-by-line, a series of one-dimensional horizontal second-order least-squares fits.

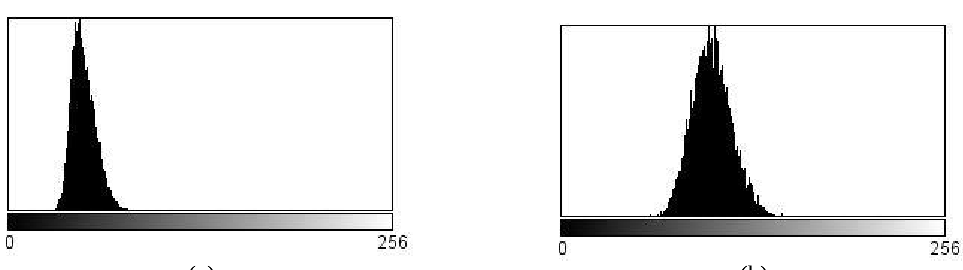
When digital data is not readily available, analog output is usable, but can be problematic for a variety of reasons. The first, of course, is re-digitizing the analog signal to acquire a cube of numbers. Depending on the array, the analog output may or may not match the full array output. For example, 640x512 arrays typically crop off the top 16 and bottom 16 lines, and output 480 lines of analog RS-170, limiting the ability to perform full-frame 3D-Noise measurements. Another common issue is 320x240 arrays are usually pixel replicated to 480 lines; however, this does not strictly alter the statistics of the result. Getting the horizontal pixels properly sampled and well-registered can be difficult, but if there is a "phase slip" between the real pixels and the digitizing, the impact on the value of the final standard deviations should be small; the greater concern would be if the analog output is poor, and the noise is filtered horizontally in the digital-to-analog conversion. This may be noticeable by having an analog horizontal MTF measurement notably lower than the vertical MTF, if that measurement is made, or by obvious and strong horizontal blurring of laboratory targets, especially if a 4-bar target is clearly resolvable through the system display, but unresolvable on a high-resolution monitor of the analog output.

Even for good analog outputs with perfectly registered digitization, the gain and level of the sensor are often problematic, most especially for systems with Automatic Gain Control (AGC, also called Automatic Gain and Level, or AGL). If the system is auto-leveling, the SITF measurement cannot be performed; no matter what temperature the blackbody is set to, the average grey level will remain the same. Note that having a "level knob" doesn't necessarily mean the system is not auto-leveling. It is common that the level knob sets the *average grey level of the scene*, but the system is still adjusting the digital values to meet that mean value; this is really more of a "brightness" control than a "level."

For this and other reasons, it is prudent to assume the worst when working with analog data, and expect that the gain and level will change if the scene is changed. Therefore, one is required to use a "target-in-scene" method for performing the 3D-Noise measurement. In this measurement, a square target whose differential temperature ΔT (with respect to the background temperature) is known is in the scene, say $\Delta T = 2.0$ K. The target should cover a reasonably large number of pixels; a good rule of thumb is to be at least 10% of the vertical FOV. An edge target (usually with a clear aperture that looks like a semicircle) will also do, with an equivalent selectable area. The idea is to be able to select an area within the target of uniform signal, sufficiently large that the average over the selected area (and over all frames) is large enough that the zero-mean noise will average to zero, so that the resultant signal value S_2 accurately represents the grey-level of the target temperature T_2. A "clean" area of the scene is selected as the background, of temperature T_2 ($T_2 - T_1 = \Delta T$), which will have a similarly averaged signal S_1 ($S_2 - S_1 \equiv \Delta S$). This area should be large enough to accurately compute the noise statistics, that is, at least 100x100 pixels. With the sensor gain and level held constant, and the temperature of the target stable, the gain and level are locked in place by the scene, even for sensors with AGC. The calibration of the image frames is then the ratio of ΔS and ΔT, which is applicable only when the noise is measured in the same noise cube.

Making an accurate calibration with analog signals requires judicious use of gain, level, and target temperature. All things being equal, a larger target temperature is better. On all blackbody system, the reference temperature T_1 is found via an external sensor, such as a thermocouple in direct thermal contact with the target wheel. If there is any discrepancy between the thermocouple reading T_1 and the actual apparent emissive temperature of the scene background, this error will be minimized by making ΔT large, in addition to minimizing other errors in offset of the target or the background, such as residual nonuniformity. However, the use of larger target temperatures will usually require using a lower gain, so that the background is not clipped, and the target does not saturate. For fully AGC systems, this may not be an option. For systems with manual gain, too high a target temperature means too low a gain, such that the binning becomes too coarse, a problem discussed above. If the binning becomes too coarse, the manual gain must be increased, and the level adjusted such that the background is not clipped; if after these adjustments, the target is saturating, then the target temperature must be reduced. The goal, then, is to have the gain set as low as possible while still being high enough to well-sample the values of the noise; in practice, this is usually a medium-to-high gain setting. Simultaneously, the level must be set high enough for the background not to be clipped, while the target temperature should be as large as possible without saturating.

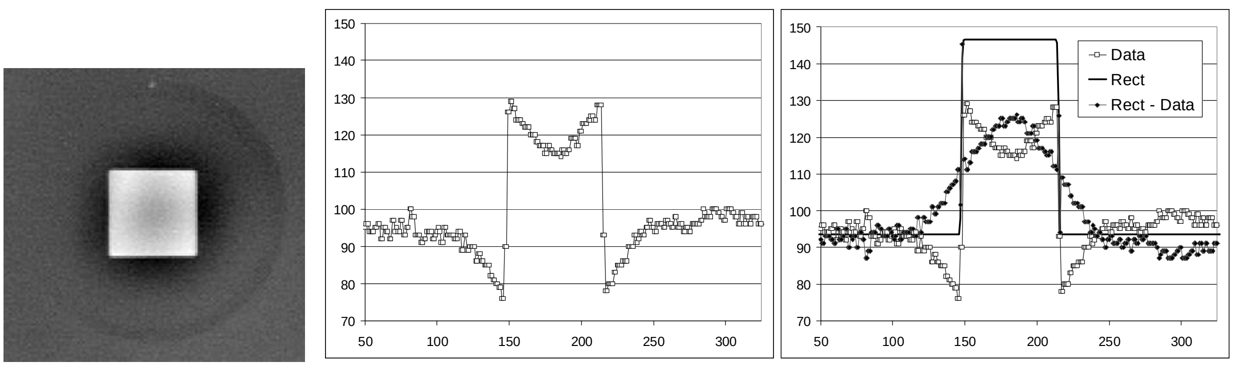
However, beyond saturation, one also needs to insure that the gain is linear in the region of interest. This can be especially difficult in fully AGC systems, but a histogram analysis can give indication if the gain is near the nonlinear region. While it is necessary, it is not sufficient that all the background pixel values be greater than zero and all the target pixel values be less than 255 (for 8 bit digitization). If the gain curve is nonlinear near the extremes of the top of bottom of the displayed temperature range, this will increase the useful dynamic range, but can skew the values of the target and/or background. Visual inspection of the image histogram will show this effect by having skewed pixel distributions that appear mostly Gaussian towards the median grey values, but drop more steeply toward the edges (0 and 255), such as in Fig. 4a. If this is the case for the background, the level must be raised (Fig 4b); if for the target, either the target temperature must be lowered or the gain reduced. In AGC systems, one is usually limited to altering the target temperature, and hoping the histogram behaves. It may be necessary to have another, hotter object in the scene (such as a human hand), even of limited spatial extent (just at the corner, perhaps) to force the AGC to reduce the gain so that the target is not saturated, and is in the linear response region of the gain. Of course, the gain-forcing hot object may well saturate the sensor at its location, but not being used as the target, its histogram is unimportant.



(a) (b)

Figure 4. (a) A "suspicious" histogram for background values. The deviation from the Gaussian distribution indicates the possibility that the lower end is in a nonlinear gain region of the AGC. (b) Raising the level recovers a more Gaussian noise distribution.

Other situations require even more clever adjustments. A good example is trying to make a calibrated 3D-Noise measurement of a Barium-Strontium-Titanate (BST) uncooled pyroelectric sensor with a diffuse chopper. To keep the dynamic range of the scene roughly constant while presenting a scene that changes with time (pyroelectric sensors only respond to changes), a diffuse chopper is used, such that each frame is actually the difference between the imaged scene and a highly blurred version of the scene. This causes a hardware-level edge enhancement, which while improving the sensor's dynamic range, makes calibration more difficult (Fig 5a). A cross-section of the target shows sharp edges, but a bowed top (Fig 5b). There is no area selectable that corresponds to any known temperature calibration. An understanding of the underlying physics allows one to extract the correct answer, however. The target edges are still clearly defined, and fitting a rectilinear (rect) function to this width and subtracting the cross-section results in a curve that is obviously the blurred version of the target, but only after the height of the rect has been properly adjusted (Fig 5c). If the rect is higher or lower, the result of the subtraction will have an obvious discontinuity, so either by inspection or by computer algorithm, the height of the "original" target can be recreated such that the blurred target is also reconstructed after the subtraction. More rigorously, one could fit a function that is the difference between a rect (matching the width of the edge-enhanced target) and that same rect convolved with a Gaussian, and even simultaneously recover the Gaussian blurring function. In either case, the retrieved rect height in greyshades corresponds to the target-background difference in temperature, and can be used to calibrate the 3D-Noise calculation. It is of course prudent to choose an area of background that is a subset of a sufficiently large ambient background such that none of the selected area contains blurred elements of area either hotter or colder than ambient.



(a) (b) (c)

Figure 5. (a) Square target with unavoidable hardware edge enhancement. (b) The cross-section shows that the calibration of this target is not directly possible. (c) Understanding the underlying physics, however, allows one to extrapolate the correct target value.

4. 3D NOISE IN PERFORMANCE PREDICTION MODELING

The ultimate utility of the 3D-Noise methodology and the laboratory measurement is in characterizing the performance of the measured sensor. At some level, a model must be applied. Even if there is a simple metric, (e.g. a pass/fail specification requirement on the level of total spatial noise) that metric is only useful or accurate if it was derived from a performance prediction model that indicated that the specification threshold corresponded to a performance threshold (e.g. a level of spatial noise that makes the probability of identification out of a specific target set, at a specified range, drop to below 50%.). The standard FLIR performance model, NVTherm (developed and continuously updated at NVESD), typically uses 4 components of 3D Noise: σ_{tvh}, σ_{vh}, σ_v, and σ_h; that is, the random spatio-temporal noise and the fixed-pattern noises. Efforts are underway to implement the proper impact of the remaining terms, which typically have been small and/or associated with "bad sensors," but may need to be considered in light of uncooled FLIRs, which are lighter, smaller, and cheaper, but also typically noisier (often, noisier in inverse correlation to the smallness, lightness, and cheapness of the sensor). From a modeling perspective, the important distinction between temporally varying noise and fixed-pattern noise is that the eye will integrate the temporally-varying noise to some degree, mitigating its impact, where fixed-pattern noise is obnoxiously and unavoidably present at all times. As such, even when the spatial noise terms are less than σ_{tvh}, it is important to include these spatial noises in modeling performance, since they affect performance more than the temporal components. Care must also be taken to insure that each type of noise is measured under the conditions that the sensor will be used, and the model is trying to predict. If typical operation and performance specifications are based upon any frame integration, the sensor must be in frame integration mode when the measurement is made. For uncooled sensors, fixed pattern noise typically increases with time[4], between NUCs, and the measurement and model must agree to use the "best case" (right after a NUC), "worst case" (usually, just before a NUC, minutes later) or some intelligent aggregate (a solution beyond the scope of this paper). The f-number and transmission of the optics in the system (or measurement setup, for component FPAs) need to be recorded with the measurement, so when the noise values are input into NVTherm, they can be correctly weighted to the system conditions; this can be easy to miss if the modeling is done months later and miles away by a different person than the one who made the 3D-Noise measurement.

5. CONCLUSIONS

Accurate laboratory characterization of infrared sensors is important for the entire life cycle of sensor design and testing. Since the two most fundamental sensor requirements are resolution and sensitivity, MTF and noise are the most important measurements. For noise, modern FLIRs (especially focal plane arrays) are almost always poorly characterized with a simple NETD, and require the use of the full 3D-Noise methodology. However, to be of proper use, 3D Noise must be measured correctly, calibrated accurately, reduced properly, and finally, modeled appropriately. This requires at least some understanding of the underlying algorithm, the data processing, how it will be used, and in some cases, the underlying physics and engineering of the sensor. It is the responsibility of any FLIR measurements laboratory to understand these elements and to apply them.

ACKNOWLEDGEMENTS

The authors would like to thank both Eddie Jacobs and John Horger for useful discussions.

REFERENCES

[1] J.A. D'Agostino, C.M. Webb "Three-dimensional analysis framework and measurement methodology for imaging system noise," Proc. SPIE Vol 1488, p 110-121 (1991).

[2] R. Vollmerhausen and R. Driggers, "Modeling the target acquisition performance of staring array imagers," Proc. IRIS Passive Sensors, vol. I, 211-224, (1998).

[3] P. Bijl, M. A. Hogervorst and A. Toet, "Identification of military targets and simple laboratory test patterns in band-limited noise," SPIE Proceedings of the Defense and Security Symposium, Orlando, FL, March 2004.

[4] S. Sousk, P. O'Shea, V. Hodgkin, "Measurement of uncooled thermal imager noise," SPIE Proceedings of the Defense and Security Symposium, Orlando, FL, March 2005.