

User's Guide for Noise Level

(Versions 3.2 & 3.3)

By Robert Lay

19 December 2005

Document Revision 21 December 2005

References

- [1] <http://www.dpreview.com/reviews/canong6/page14.asp>
- [2] http://www.imatest.com/docs/tour_q13.html
- [3] <http://www.imatest.com/docs/noise.html>

Introduction

This new version of the Noise Level Program is designed to provide a measurement of the noise level in a test image. The noise level is measured on an image of an 18% Gray Card and is expressed as the standard deviation of the digital brightness level of the pixels in a central subset of the virtual image.

The conditions under which noise is measured must be standardized in order that the measurements have value for comparison purposes. This program requires that the measurement be based on the Gray Card image.

This program is not intended for use in the measurement of noise in a general photographic scene.

If the usage requirements for this program are followed, the program will serve as a secondary standard for the purposes of comparing cameras or for comparing results from a given camera with different settings, such as the ISO speed setting and exposure duration.

Principles of Operation

The basic principle of operation is based on the concept of "standard deviation". As the value of each pixel in the image is read from the virtual image file (JPG, BMP, GIF, ICO), the values of luminance in each of the three channels, R, G, and B are each processed to obtain values for the mean, the variance and the standard deviation of the values of the pixels in the array of pixels for each of the RGB channels.

In addition, the *combined RGB* luminance of the pixel, based on the average luminance of the 3 channels for that pixel, is also processed to obtain a value for the mean, variance and standard deviation of the values of the pixels in the array of pixels for the *combined RGB* channels.

In order for the standard deviation to be representative of noise, this process assumes that the image is of a standard, 18% reflectance Gray card. Ideally, the pixels in that image will be uniformly illuminated and there will be no variation in the combined luminance values or in any of the three primary color channels other than that caused by noise. However, in a typical image taken from a Gray card there will be some

variation of the illumination over the height and width of the image and some variation in the sensitivities of the three channels at each pixel, and those variation would also be considered as noise.

Conversely, if only a small area at the center of the image is sampled, it is less likely that slow variations in brightness would affect the measured values of noise, which is assumed to be a random process manifesting itself independently in each pixel. To that end the program samples a rectangular area at the center of the image whose height is 9 pixels and whose width is 12 pixels. The primary reason for restricting the size of the array of pixels is to reduce the possibility of anything other than noise from entering into the calculation of standard deviation. For example, the larger the sample array, the more likely it is that the brightness values of the pixels will decrease with distance from the center due to vignetting.

The standard deviation of a set of N numbers X_1, X_2, \dots, X_N is denoted by s and is defined by

$$s = \sqrt{\frac{\sum_{j=1}^N (X_j - \bar{X})^2}{N}}$$

NoiseLevel implements the above definition by sampling 108 pixels in an array of 9 x 12 pixels from the center of the user supplied test image. Since each pixel is represented by a triad of 3 values for Red, Green and Blue channels, the standard deviation for each channel is calculated and displayed separately. The mean value, \bar{X} , and the variance, s^2 , for each channel are also displayed.

In order to obtain the standard deviation for the luminance, the three primary channel values for each pixel are first averaged to obtain a luminance value for that pixel. The resulting array of 108 luminance values is then processed for standard deviation exactly as the individual channels were processed and that is displayed as the combined value of standard deviation (for the combined RGB) along with the combined mean and combined variance values. Unless there is some anomaly in the camera sensor, the four different values of standard deviation will be close to one another and the combined RGB values will be generally less than some of the individual channels. It is not unusual for a given camera to exhibit more noise in one channel than in another. In other words, the noise is not necessarily balanced across the three channels, except when the image is B&W.

The method used by this program can be compared with that used in the noise tests conducted by *dpreview.com*. A specific example of such a test, which shows data taken from Canon models G5 and G6 as well as other cameras, will be found at [1].

Data taken with **NoiseLevel** version 3.3 and a Canon Powershot G5 at ISO settings of 50, 100, 200 and 400 compared favorably with data presented by *dpreview.com* in the above reference.

Simple Mathematical Example

The internal workings of **NoiseLevel** can be better understood by means of a simple numerical example using real pixel data. The major simplifications that we will use are that this example will use data from of a "Lens Cap On" test, rather than a Gray Card test.

The table below is a representation of a sample 5 pixel by 4 pixel image. Each cell of the table represents a pixel and the value in that cell represents the luminance of that pixel. The value is shown as one number,

but that one number represents the average of the three channels, Red, Green, and Blue. Note that the values are relatively low in comparison with the full range of possible values of 0 to 255. These values were intentionally chosen as representative of actual noise levels in a digital camera. Also, keep in mind that 0 represents pure black and 255 represents pure white.

1	2	3	1	5
0	0	0	1	0
1	2	3	0	6
5	3	3	1	4

The Histogram

The program begins by searching through the table and counting the occurrences of value = 0, which in this example is 5 occurrences (or a frequency of 5). It stores that in the histogram array at histogram cell #0.

The program then repeats that operation to determine the number of occurrences of value = 1. It stores that count in histogram cell #1.

This process is repeated for each possible value from 0 through 255, and at the completion of that process the histogram array contains in each cell the number of occurrences or “frequency” of each brightness value from 0 to 255. In this example the histogram array values, starting at histogram cell #0 are 5, 5, 2, 4, 1, 2, 1, 0, 0, 0..... All histogram cells from 7 through 255 contain 0, because there are no brightness values above 6 in this sample image. The resulting histogram is plotted in the upper left portion of the window.

Calculation of the Standard Deviation

The program next computes the mean value, which is nothing more than the sum of the 20 values shown (41) divided by the number of pixels (20), which yields 2.05.

Next, the program calculates, for each of the 20 pixels, the difference between that pixel value and the mean, as follows:

```

-1.05  1.1025
-0.05  0.0025
0.95   0.9025
-1.05  1.1025
2.95   8.7025
-2.05  4.2025
-2.05  4.2025
-2.05  4.2025
-1.05  1.1025
-2.05  4.2025
-1.05  1.1025
-0.05  0.0025
0.95   0.9025
-2.05  4.2025
3.95   15.6025

```

2.95	8.7025
0.95	0.9025
0.95	0.9025
-1.05	1.1025
1.95	3.8025

The second column shows the subsequent step, which is the square of the deviation. The squares of the deviation are then summed (55.1875) and divided by the number of pixels (20), and the final result, known as the “variance”, is 2.759375 .

The square root of the variance is the “standard deviation” (1.66), also known as the “root mean square (rms) deviation”.

In summary, for this example, we can say that the rms value of the deviation numbers in the sampled pixel array is 1.66 units of digital brightness (on the scale of 0 to 255). It is also very important to recognize that in a “Lens Cap On” test, the values found in each pixel are relatively low compared to other test images. It is not unusual to find as few as 20% of the pixels in the image deviating due to noise. Remember that in the deepest shadow areas where the pixels have values between 0 and 5, very few pixels will have a deviation from the mean of more than 1 or 2 units (on the scale of 0 to 255). As will be seen later, the deviations that we can expect to see when analyzing images of Black, Gray or White Cards, the deviations measured in a “Lens Cap On” test, for a given camera, ISO setting and exposure time, will be considerably lower levels than those found in any of the other tests. That is one of the reasons why the author of this program does not recommend a “Lens Cap On” test or a Black Card Test or a White Card Test for purposes of making comparison tests. This subject will be discussed further in a later section.

Step-By-Step Procedures for Using the Program

The Step-by-Step procedure for taking the test image and analyzing it follows.

Step 1

Set up a standard Gray card in an environment of diffuse lighting. Avoid harsh lighting as that will tend to produce hot spots or shadows in the image. The most uniform lighting possible is best. The lighting does not have to be at a particular color temperature, but it should be nominally white light from either daylight or tungsten source.

Step 2

Set up the camera to take the pictures so as to facilitate a uniform approach from shot to shot. You will probably be interested in measurements over a wide range of shutter speeds and ISO settings, so a tripod will be an advantage, although not absolutely necessary. Ideally each frame of test shots should have the image completely filled by the Gray card. Either use a sufficiently large gray card or zoom in to fill the frame. However, since a small array of pixels from the central area is used for the analysis, filling the frame completely is not essential.

The camera can be set to B&W mode, if desired, but the results may be slightly different than when using full color. There are two reasons for that – one is that the algorithms used in the camera for B&W mode may provide different gamma encoding than for color, and the other is that we do not want to require that the Gray Card be precisely neutral nor do we want to insist that the lighting be a specific color temperature.

We do not recommend any Black or White Card or “Lens Cap On” substitutions. There are very specific reasons for standardizing on the Gray Card test, which will be discussed further in a later section.

It is strongly recommended that JPG files direct from the camera be used rather than RAW files that have been processed to JPG. The reasons for this will be discussed at length in a later section.

Step3

With each shot, manually record the shutter speed and ISO settings of the camera or make sure that the EXIF data will be preserved for later reference. You will probably want to plot curves of noise vs. ISO settings or curves of noise versus exposure times in order to see the correlation graphically. Take the exposure using the maximum size image provided by the camera. This is necessary due to the fact that when smaller size images are used the camera actually down-samples the larger image in order to create the smaller image, thus producing an image with much lower noise levels than those from the larger image.

Step 4

Once the entire battery of test shots is complete, process your images to computer files in the usual way but with the following suggestions:

Make note of any noise reduction software being applied during normal processing. Ideally, such noise reduction software should be turned off during the initial cataloging of the images. For example, Adobe Camera RAW automatically applies a specific level of noise reduction under the “Detail” tab unless you take the trouble to turn that off.

Make the final save as Mode 8 bits in either BMP, JPG or GIF format.

Step 5

Using **Noise Level**, open each image file. Record the standard deviation values shown for each of the 4 columns – Red, Green, Blue and Combined. The Combined column is the one you will normally use for all parametric studies. However, the primary color channels are reported individually in order for you to determine that there are no gross anomalies in the area sampled. Likewise, although the mean and variance data will normally be of little interest, it is provided as backup information.

Interpretation of Results

This section describes the differences one might encounter as a result of comparisons between different size image sensors, different ISO settings, different exposure times, different gamma encoding, and differences caused by down-sampling or compression. There are also subsections covering White Cards and Black Cards.

If you follow the Step-by-step instructions for the Gray Card test using **NoiseLevel**, there is no need to be concerned about other types of images that might be loaded into the program for analysis. Nonetheless, we know that there will be questions about such possibilities, and we also appreciate that it may help the user to better understand the overall process if he is familiar with all aspects of the problem. So, first we should look at what to expect from the Gray Card test, and then we will briefly describe what can be expected in other tests, even though the program was not designed to provide any useful information from such tests.

The first thing to understand is that the results of the test are embodied in one number, the standard deviation for the combined RGB. All of the other displayed data is just there in the interest of providing anything and everything of interest.

The second most important aspect of interpreting the results is to realize that the standard deviation is a linear measure of the noise – not logarithmic. So, numbers that look like they are close together *are* close together. Also, the bigger the standard deviation, the bigger is the noise. When comparing the results with other cameras or other settings, remember that the displayed results are like Golf scores – you want the lowest possible values from your camera.

Expect to see measured results larger with higher ISO settings or higher exposure times. Exposure times less than 1 second are probably not going to follow that rule, but exposure times greater than 1 second might cause significant increases in standard deviation. This is because the noise mechanism in the camera is one which is integrated over time. Therefore for exposures of 1 second or longer, that effect is likely to cause an increase in measured values.

Do not expect aperture settings to affect the standard deviation There is reason to believe that elevated temperatures can increase the noise level. Take nothing for granted. If you want to test your camera to see if temperature is a factor, it is easy enough to run a parametric analysis over a range of temperatures, say at freezing, room temperature and 105 degrees Fahrenheit. Just repeat the Gray Card test at those three temperatures with a single configuration of ISO and shutter speed, and then you will know.

Expect the standard deviation measurements to be inversely related to sensor array dimensions. The bigger the sensor array, the lower the standard deviation. Likewise, for a given sensor array size, the size of the individual sensor sites is a factor that is also inversely related to standard deviation. So, if the individual sensor sites are bigger (for a given sensor array size), the standard deviation should be lower. You might well ask how the sensor sites can vary in a given sensor array. The sensor sites each sit in the middle of a bunch of electronics. It's all a question of how much of the available sensor site area is actually available for the sensor and how much is occupied by the ancillary electronics.

Expect the standard deviation measurements to be a function of down-sampling. As a given image is down-sampled to a smaller virtual image size the standard deviation decreases almost linearly. I can think of no reason to conduct the noise test on a down-sampled image. However, be advised that the results of such a test will not be comparable with images that were not down-sampled.

At this time we do not have enough data from users of the program to determine whether or not there is any correlation between the size of the virtual image and the measured standard deviation. There should not be any correlation, because the program is only looking at a small array of 9 pixels by 12 pixels and doesn't even know the size of the full image.

The results of a "Lens Cap On" test have already been discussed in a previous section.

Results from White Cards or Black Cards

This subsection describes the White Card test and Black Card test – each of which will differ from the Gray Card in certain very specific ways.

A Black Card shot can be taken in three ways – it can be exposed to place the black card image at its proper level, which is at a value characterized as "maximum black with detail" (brightness levels of 5 to 15), or it can be given a "normal" exposure, which places it in the neighborhood of middle gray (127). We

will discuss the results of the former placement only, since the latter placement is a departure from common and proper photographic procedure.

A properly exposed Black Card will therefore have a mean value of between 5 and 15. Due to the normal gamma encoding applied to the image in the camera when the camera is configured for JPG formatting, the standard deviation, or rms deviation due to noise will be found to be significantly lower than in a Gray Card test. We believe this to be a result of the lack of expansion of detail in the shadows during gamma encoding when the camera is configured for JPG formatting. In other words, the rms deviation due to noise should be normalized by the derivative of the range of pixel values with respect to the range of f-stops corresponding to those pixel values in the shadow end of the transfer characteristic curve.

A White Card shot can be taken in two ways – it can be exposed to place the white card image at its proper level, which is at “maximum white with detail” (brightness levels of 240 to 255), or it can be given a “normal” exposure, which places it in the neighborhood of middle gray (127). We will discuss the results of the former placement only, since the latter placement is a departure from common and proper photographic procedure.

A properly exposed White Card will therefore have a mean value of between 240 and 255. Due to the gamma encoding applied to the image in the camera when the camera is configured for JPG formatting, the standard deviation, or rms deviation due to noise will generally be measured at a level higher than that measured in a “Lens Cap On” test and lower than that measured in a Standard Gray Card test. We believe this to be a result of the fact that the mid tones are given more expansion (higher contrast) during gamma encoding, while the highlights are actually compressed to a much lower range of steps. In other words, the rms deviation due to noise should be normalized by the derivative of the range of pixel values with respect to the range of f-stops corresponding to those pixel values in the shadow end of the transfer characteristic curve.

Results from General Photographic Images

This subsection describes the situation of a test image using a typical photograph.

Consider an image that has an extreme of either information or noise. It would have a mean value of about 127 and would have each pixel deviating from the mean by 127, maximum. If you plug those values into the equation for the standard deviation, the result is 127. In other words, the ratio of the standard deviation to the mean gives you an indicator that I call the entropy of the image, and it will not exceed unity. Images that we have measured just to get a feel for typical values have values of about one third of the mean value. In other words, a typical photograph will measure about 30 to 40 in the NoiseLevel program.

However, that is in no way a measure of the noise in that image, because a computation of standard deviation of the brightness levels in the pixels of an image has no way of distinguishing between information and noise.

An image with very little information and large areas of one color and luminance will give relatively low results in **NoiseLevel** – approaching the noise level as a lower limit. In other words, a photograph that has little or no information will actually be revealing its noise level, because there is nothing but noise.

Further Considerations on the Importance of the Gray Card as the Test Image

As has been mentioned in earlier sections, there is a basis for insisting that the program be used with an 18% Gray Card instead of other possible test images, such as the “Lens Cap On” test, the Black Card test, the White Card test, or a test using an ordinary photograph.

The ordinary photograph contains useful information about the scene. That information is seen only as noise by any analysis that measures the standard deviation across a sample of its image data. In other words, a measurement of the standard deviation in a typical photograph does include the effects of noise, but the dominant factor in determining the standard deviation is the information in the scene – not the noise. Therefore, we reject the typical photograph as being unsuited for our purposes.

Only one of the remaining possible test images closely approximates the average brightness of the average photographic scene, and that is the Gray Card. Only one of the remaining three test images can be used if we are to have a universal standard for comparisons.

We reject the “Lens Cap On” test for the following reasons:

- It is very easy to have a light leak that completely destroys the validity of the test.
- It is dependent upon the shutter duration.
- It does not reveal the effect of the camera’s gamma encoding.

That leaves only the White Card and the Gray Card as the only viable candidates. We know a little about how Imatest measures noise from [2] and [3]. We know the algorithm used to calculate rms deviation is the same as we use in **NoiseLevel**, as described in an earlier section. We also know that they plot the noise measurement at each point of the transfer characteristic using a 20-step step chart image of either 0.1 or 0.15 log density steps. Their data is sometimes displayed as a percentage and sometimes displayed in f-stops. In both cases the scale used for noise is overly complex and very dependant upon the use of the step chart. Another interesting aspect of the way Imatest publishes their noise data is that they normalize their rms deviation data by dividing it by the derivative of the range of pixel values with respect to the range of f-stops corresponding to those pixel values in the relevant area of the transfer characteristic curve. For a Gray Card image, that would mean normalizing their rms deviation results by dividing by a derivative measured in the middle gray area of the transfer characteristic curve. In other words, all three test images should actually be normalized in order that they indicate the actual noise in terms of pixel values. We have not done that for the simple reason that our concept of a tool for measurement of noise should be simple, free, and should not require the use of any precision standards, such as a Kodak step chart – nor should it require the careful measurement of gradients throughout the transfer characteristic curve.

In our opinion, the Gray Card image is the only practical alternative to the considerably more expensive, complicated and automated system using the step chart.

Therefore, unless someone can convince us that the White Card would be a better choice, we continue to recommend that **NoiseLevel** be used with a Gray Card image.

RAW Files Not Recommended

It is recommended that test images taken for use with **NoiseLevel** be processed to JPG in the camera and loaded into **NoiseLevel** in the form in which they are delivered from the camera and without

modification. To do otherwise would seriously undermine the entire noise measurement process from the standpoint of standardization.

Users should certainly experiment with RAW files in whatever ways they see as beneficial to their interests and objectives, whatever they may be. However, the use of **NoiseLevel** as a tool for standardized reporting of noise levels for comparison purposes requires uniformity in the acquisition of the image files. It is impractical to expect that RAW processing would be done the same by all users – especially since there are numerous RAW processing programs currently in use, no two which are alike.

Different cameras (especially between different manufacturers) may process to JPG differently from one to another, but within a given make and model the processing to JPG is consistent across all users.

Of primary interest is that a given user be able to compare the performance of his camera against the entire community of users of like make and model. Of secondary importance is that users with different makes and models be able to make comparisons between makes and models for the purpose of determining whether or not a given make or model has significant advantages or disadvantages with respect to the various makes and models in a given class (P&S, Prosumer, Professional, etc.). Also of secondary importance is the educational process which a user experiences when testing his camera for its noise performance against various parameters, such as temperature, exposure duration, and ISO setting. Not only does the user acquire useful data, but he also gains a better understanding of the origins of noise and how it affects his images.

All of these objectives and benefits to the user community are extremely dependant upon a consistent and scientific approach to the problem, which requires standardization of all of the process steps.