Quantitative evaluation of the signal-to-noise ratio and of its improvement (or degradation) by digital filters

N. Bonnet, S. Lebonvallet, H. El Hila, G. Colliot, A. Beorchia

To cite this version:


HAL Id: jpa-00248662
https://hal.archives-ouvertes.fr/jpa-00248662
Submitted on 1 Jan 1991

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Quantitative evaluation of the signal-to-noise ratio and of its improvement (or degradation) by digital filters

N. Bonnet, S. Lebonvallet, H. El Hila, G. Colliot (*) and A. Beorchia

Unité INSERM 314, 21, rue Clément Ader, 51100 Reims, France
(*) Société BIOCOM, 24 avenue de la Baltique, 91942 Les Ulis Cedex, France

(Received 5 November 1990, revised 1 March 1991, accepted 11 March 1991)

Résumé. — Le rapport signal/bruit d'un ensemble de données (spectre, image...) peut être évalué quantitativement si une deuxième réalisation des données peut être acquise dans des conditions identiques. Les applications possibles de ce calcul concernent : la comparaison de différentes procédures expérimentales d'acquisition des données (relativement au rapport signal/bruit) ; l'évaluation de filtres numériques en ce qui concerne l'amélioration ou la dégradation du rapport signal/bruit qui résulte de leur application au double ensemble de données.

Abstract. — The global signal-to-noise of a data set (spectrum, image...) can be quantitatively evaluated if a second version of the data can be recorded in identical conditions. The possible applications of this computation are : the comparison of different data acquisition procedures (with respect to the signal-to-noise ratio) ; the evaluation of digital filters applied to the original double data set, concerning the improvement or degradation of the signal-to-noise ratio which results from their application.

1. Introduction.

The concept of signal-to-noise ratio is involved in a large number of operations which deal with signal or image processing (acquisition, smoothing, enhancement, restoration...). However, this concept is most of the time involved from a qualitative point of view, except when noise is simulated. In real situations, where the amount of noise is unknown, very few reports give quantitative values for the signal-to-noise ratio and the effect of noise is described by very inaccurate adjectives (weak, strong...).

It seems that many physicists are convinced that the signal-to-noise ratio can be improved (or degraded) by special processes but that the absolute value of this ratio cannot be estimated precisely.

In some branches of Physics however (Electron Microscopy for instance), it has been shown that the signal-to-noise ratio can be estimated [1, 2, 3]. The requirement for this estimation is only a slight modification of the experimental procedure.

In this note, we would like to popularize this method so that other fields could take advantage of it and to show some applications. These applications concern:
a) the evaluation of a data acquisition procedure with regard to the signal-to-noise ratio (in order to allow an objective comparison of different variants of the procedure). This task will be illustrated through digital image acquisition in Scanning Electron Microscopy (SEM);

b) the evaluation of digital filters, again from the point of view of the signal-to-noise ratio.

The present paper is organized as follows:

In section 2, we give a method for the quantitative evaluation of the signal-to-noise ratio.

In section 3, we demonstrate by a simulation the validity of the procedure and we give an example of application of this method for the acquisition of digital data in Scanning Electron Microscopy.

In section 4, we give examples of application to the evaluation of different digital filters.

2. Outline of the method for the evaluation of the signal-to-noise ratio.

We concentrate on the case of a signal $s$ degraded by an uncorrelated additive noise with zero average:

$$i(x) = s(x) + n(x)$$

where each quantity is a function of a given variable $x$ (time, energy, spatial coordinates...).

We define the quadratic signal-to-noise ratio (SNR) as the ratio of the signal variance to the noise variance:

$$\text{SNR} = \sigma_s^2/\sigma_n^2.$$  \hspace{1cm} (2)

It is clear that one experimental realisation of the process described by equation (1) cannot lead to any estimation of SNR, since there is only one equation for two unknowns.

One way to evaluate SNR is to start with two equations. This is possible if one is able to get two successive realisations of the degraded signal $i$, in identical conditions:

$$i_1 = s + n_1$$

$$i_2 = s + n_2.$$  \hspace{1cm} (3)

In this context, what will be called "signal" ($s$) is the part of the measured quantity which is common to both realisations (deterministic part) and what will be called "noise" ($n_1$ and $n_2$) is the changing (or stochastic) part.

One could use the mean of the two realisations as an estimation of the signal and the differences between the individual realisations and the mean as two estimations of the noise. This approach would yield an estimation of the signal-to-noise ratio as:

$$\text{SNR} = \frac{\sigma_s^2}{\sigma_{\Delta i_1}^2 + \sigma_{\Delta i_2}^2}.$$  \hspace{1cm} (4)

where

$$\bar{i} = (i_1 + i_2)/2$$

$$\Delta i_1 = i_1 - \bar{i}$$

$$\Delta i_2 = i_2 - \bar{i}.$$  

But, as it was described in the context of multiple image averaging [4] this estimation cannot be very efficient because $\bar{i}$ is itself noisy, so that the variances in the denominator of (4) are underestimated and the SNR is overestimated, especially for low SNR. This will be demonstrated through a simulation in section 3.
Assuming that the two realisations of the noise \((n_1\) and \(n_2\)) are uncorrelated, the signal-to-noise ratio can be deduced more efficiently from the cross-correlation function between the two experimental realisations \(i_1\) and \(i_2\).

The cross-correlation function is defined as:

\[
C(x') = i_1(x) \otimes i_2(x)
\]

\[
= \frac{1}{N} \sum_{x=1}^{N} i_1(x) i_2(x + x')
\]

\[
= \frac{1}{N} \sum_{x} [s(x) + n_1(x)] [s(x + x') + n_2(x + x')]
\]

\[
= s(x) \otimes s(x) + s(x) \otimes n_2(x)
\]

\[
+ n_1(x) \otimes s(x) + n_1(x) \otimes n_2(x)\]

where \(N\) is the number of sampling points used for the summation.

The three last terms in the right part of the equation vanish because the noise is uncorrelated with the signal and the two realisations of the noise are uncorrelated. So, the cross-correlation function between \(i_1\) and \(i_2\) reduces to the autocorrelation of \(s\), i.e. \(s(x) \otimes s(x)\).

So, we can conclude that the maximum value of \(C(x')\) is the signal energy \(s^2\) and a simple relationship exists between the height of the cross correlation peak and the signal-to-noise ratio.

A slight modification of the cross-correlation function \(C(x')\) leads to the cross-correlation coefficient \(\rho(x')\) which always has values between \(-1\) and \(1\):

\[
\rho(x') = \frac{C(x') - \bar{I}_1 \cdot \bar{I}_2}{\sigma_1 \sigma_2}
\]

where \(\bar{I}_1\) and \(\bar{I}_2\), \(\sigma_1\), \(\sigma_2\) are the means and standard deviations of signals \(i_1\) and \(i_2\) respectively.

Now, it can be shown easily that \(\rho_m\), the maximum value of \(\rho(x')\), allows to estimate \(\sigma_s^2/\sigma_n^2\) and that the signal-to-noise ratio described by equations (2) can be expressed as:

\[
\text{SNR} = \rho_m/(1 - \rho_m).
\]

In the general case, the cross-correlation function must be computed as a function of the variable \(x'\) and much more efficient computations are performed through Fourier transforms of the experimental data [5]:

\[
C(x') = \sum_{x} i_1(x) i_2(x + x')
\]

\[
I_1(k) = F.T. [i_1(x)]
\]

\[
I_2(k) = F.T. [i_2(x)]
\]

\[
C(x') = F.T.^{-1}[I_1(k) I_2^*(k)]
\]

where \(F.T.\) and \(F.T.^{-1}\) are the direct and inverse Fourier transformations, respectively, and \(+\) states for the complex conjugate.

However, if the experimental set-up is such that the two digitized signals \(i_1\) and \(i_2\) can be considered as correctly registered, the maximum value \(\rho_m\) of \(\rho(x')\) will be obtained...
for \( x' = 0 \). In that case, the cross-correlation coefficient can be obtained more easily by applying the definition (5) for \( x' = 0 \):

\[
\rho_m = \rho (0) = \frac{i_1(x), i_2(x) - \bar{i}_1 \cdot \bar{i}_2}{\sigma_1 \sigma_2}
\]  

(9)

3. Application 1: quantifying the improvement of the signal-to-noise ratio by time integration.

In many situations, where the signal is obtained by counting physical events, the origin of noise can be found in the statistical fluctuations of the counting process (quantum noise). It is well known that the only way to improve the signal-to-noise ratio in that case is to increase the number of physical events which are counted, by increasing the counting time for instance. This is of common practice in digital image recording: when the amount of signal in a single video frame is insufficient, one is used to sum up several individual images in order to improve the signal-to-noise ratio. Simple statistical considerations show that SNR is supposed to increase in proportion to the number of frames which are added together (in proportion to its square root when the SNR is defined as an amplitude ratio instead of an energy ratio).

This simple law concerning the relative variation of SNR does not provide any indication concerning the absolute value of the SNR. Such an absolute value would however be useful because in many situations, the counting time cannot be increased indefinitely (drift of the specimen, radiation damage...) and it would be interesting to stop the counting process when a given value of the SNR is attained.

The procedure described above allows us to obtain a quantitative evaluation of the SNR. This will be illustrated with image sequences. The first sequence (Fig. 1) is a simulation: two

Fig. 1. — Aerial photograph (a) with an increasing simulated Poisson noise (b, c, d).
noisy versions of a noise free image were created by a Poissonisation simulating the quantum noise which would result of a limited number of counts in every pixel (the mean number of counts per pixel was 20, 5 and 2, giving a SNR of 5.9, 1.5 and 0.66, respectively). The results obtained from the two methods described in the previous section are compared with these theoretical values in figure 2. Obviously, the method based on the cross-correlation coefficient between the two images of each couple is superior, especially for low SNR situations.

Fig. 2. — Comparison of the SNR estimated by two different methods relying on the comparison of the experimental images with the mean image (+) and on the computation of the cross-correlation coefficient between the two realisations (■).

The second sequence (Fig. 3) is a real sequence obtained in a Scanning Electron Microscope (backscattered electrons).

At video rate (25 images/second), the number of collected electrons per pixel is so small that the image is not interpretable, due to the associated quantum noise (Fig. 3a). The digital acquisition system (BIO 500, BIOCOM, Les Ulis, France), based on an image board (MVP/AT, Matrox, Dorval, Canada), is able to sum up a large number of video frames.

Figures 3b, c, d display the results obtained by adding 10, 50 and 200 frames, respectively. The absolute values of the SNR, computed from image pairs according to the procedure described above, are displayed in figure 4. One can see that the linear increase as a function of the integration time, predicted by the theory, is fairly well verified.


The performances of a given filter are generally evaluated by two complementary approaches:
Fig. 3. — Sequence of backscattered electron images obtained in Scanning Electron Microscopy by increasing the number of video frames averaged (1, 10, 50 and 200 frames are added to obtain images a, b, c, d).

Fig. 4. — Evolution of the SNR as a function of the number of integrated frames in figure 3.
the filter is studied theoretically. Its transfer function and point spread function are defined. Its effects on a given theoretical kind of noise (white noise, shot noise...) can be investigated;

when applied to a real situation, the filter effect is ascertained by eye, the operator judging that its goal has been attained better by one filter than by another one. Even if this approach can be sufficient when the filters are self consistent, i.e. when their goal is to improve the visual aspect of an image, it can be insufficient when the filter acts as a preprocessing tool. In that case, it would be interesting to define the effect of the filter on the original image by an objective criterium.

Let us give an example: in order to smooth an image degraded by an additive noise, one has mainly the choice between a linear shift invariant filter [6] (implemented as a frequency filter in reciprocal space or as a convolution filter in the direct space) and a non linear filter such as a median filter [7]. Roughly speaking, advantages and drawbacks of both filters are known: linear filters ensure in a typical situation a good signal to noise ratio improvement but lead to a resolution loss; median filters are especially suitable to reduce some types of noise (salt-and-pepper noise for instance) and do not severely degrade the signal but they are generally less efficient than linear filters in improving the signal to noise ratio.

The question is: it is possible, in a given situation (kind of noise, value of signal-to-noise ratio...) and for a given filter, to define quantitatively some parameter allowing the operator to make an objective choice between several candidate filters.

Another example is given by the reverse situation: instead of performing noise removal (low-pass filtering), suppose that we have to perform any signal (or image) processing which relies on some kind of high-pass filtering. Differentiation, for instance, is a data process commonly used in spectroscopy as well as in image processing (edge detection...). This high-pass filtering results in an unavoidable way in a deterioration of the signal-to-noise ratio [8], because the high frequency components of noise are also increased. So again, the question is: is it possible to quantify this degradation of the SNR and among several candidate filters which are able to perform the operation of interest (differentiation, enhancement...), it is possible to make an objective choice taking into account the deterioration of the SNR?

Our purpose is not to study extensively the large number of filters which have been suggested in the field of signal or image processing but to show that the procedure described above can bring a part of the answer, and can be used for each particular situation.

We give two illustrations of this procedure.

The first one concerns the process of smoothing: among the many filters which have been devised for improving an image signal-to-noise ratio we selected three commonly used ones and we applied them to the image displayed in figure 3b. The first two are convolution filters, where the resultant image is obtained by a convolution of the original image with the convolution matrix:

\[
\frac{1}{9} \begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}
\quad \text{and} \quad
\frac{1}{16} \begin{pmatrix}
1 & 2 & 1 \\
2 & 4 & 1 \\
1 & 2 & 1
\end{pmatrix}.
\]

The third filter is a classical median filter, where the resulting pixel value is the median value computed on a 3 x 3 environment.

In order to quantify the improvement of SNR, the smoothing operation must be applied to a couple of original images, obtained in identical conditions. The SNR is computed before and after filtering and then, the ratio is computed.

For the three filters described above, which give the results displayed in figures 5b, c, d, the gain in SNR is 4.33, 3.59 and 3.12, respectively.
It should be noted that:

a) the gain in SNR is not the only parameter which has to be taken into account before adopting any filter. Its effects on the image resolution must also be considered.

For this purpose, other quality factors also deduced from the computation of the cross-correlation coefficient (before and after applying the filter) can be used [9, 2]. Others indicators have also been suggested which take into account both the improvement of the signal-to-noise ratio and the degradation of the resolution [3].

b) the result obtained in this situation (the best SNR improvement for the first filter) cannot be considered as a general result. Such an experiment must be reproduced for every new situation.

The second illustration concerns the differenciation problem and the quantification of degradation of SNR. Image edge detection is generally performed through the computation and the combination of several gradients. It is a difficult problem for noisy images, so that a lot of filters have been suggested [8].

Just in order to illustrate the potentialities of the method for comparing these filters with respect to the SNR, we applied two gradient filters to image 3b, resulting in figure 6.

The first filter, which gave the result in figures 6a and b, is the sum of an horizontal filter and of a vertical filter, each of these computing the absolute values of the difference between adjacent pixels.

The second filter (results 6c and d) computes what is called a morphological gradient [10]; that is the difference between a dilatated image and an eroded image (dilatation and erosion are basic operations of the so-called mathematical morphology [11].

Fig. 5. — Application of different smoothing filters (3 × 3 convolution filters and median filter) to a noisy image.
The degradation of the signal-to-noise ratio (compared to the original image) can be computed in the same way as the improvement by smoothing filters.

In this specific case (which cannot be generalized), the figures are 0.11 and 0.31 for the first difference gradient and for the morphological gradient.

5. Conclusion.

We have shown that the signal-to-noise ratio of an experimental data set (spectrum or image) can be quantitatively estimated provided one accepts the experimental constraint of recording a second version of the data set in identical conditions. The method based on the computation of the cross-correlation coefficient is clearly superior to the method based on the comparison of the original data to their mean when the SNR is weak.

When the couple of data can be considered as registered (no energy shift for a spectrum, no space shift for an image), the computation effort is very small since only one value of the correlation coefficient is needed, which can be obtained through one pass over the data.

This quantitative estimation of the SNR is useful not only for a characterization of the experimental data set, also for a characterization of the subsequent treatments, since computing the SNR before and after processing allows an estimation of the improvement (or degradation) of the SNR they induce.

Furthermore, this double acquisition procedure can also be used:
- for an estimation of the resolution (before and after filtering) [2];
- as a double input into specially devised smoothing filters [12].

However, it should be stressed that more sophisticated procedures [3], using the concepts of Geostatistics [13], have been suggested in order to characterize a data set (before and after
filtering) by a single parameter taking into account both the signal-to-noise ratio and the resolution.

More practice (in different fields of applications) with the very simple procedure described in this paper is necessary before drawing conclusions concerning the comparison with these optimal filtering procedures.

References