3

Methods for Blind Estimation of the Variance of Mixed Noise and Their Performance Analysis

Sergey Abramov[1], Victoria Zabrodina[1], Vladimir Lukin[1],
Benoit Vozel[2], Kacem Chehdi[2] and Jaakko Astola[3]
[1]*National Aerospace University,*
[2]*University of Rennes 1,*
[3]*Tampere University of Technology,*
[1]*Ukraine,*
[2]*France,*
[3]*Finland*

1. Introduction

Modern imaging systems provide a huge amount of images nowadays. These images are of different original quality. Some of them are practically ready for exploitation, e.g., visual inspection, object recognition, etc. Other ones need to be pre-processed as, e.g., by filtering, edge detection, segmentation, compression (Pratt, 2007; Bovik, 2000; Al-Shaykh&Mersereau, 1998), etc. In the latter case, it is desirable to know noise type and characteristics (Pratt, 2007; Elad, 2010). Such information is exploited by modern methods and algorithms of image denoising (Elad, 2010; Sendur&Selesnick, 2002; Donoho, 1995; Mallat, 1998), edge detection (Pratt, 2007; Touzi, 2002) for setting proper thresholds that depend on noise statistics.

In some practical situations, noise type and basic characteristics are known in advance. An example is radar imaging by synthetic aperture radar (SAR) with known number of looks and image forming mode (Oliver&Quegan, 2004). However, there are quite many practical situations where noise type and/or characteristics are not known in advance. Images acquired by digital cameras can serve as an example where noise properties are determined by camera settings, illumination conditions (Liu et al., 2008; Foi et al., 2007), etc. Then, noise characteristics are to be estimated for each particular image subject to further processing, for example, filtering or compression (Liu et al., 2008; Foi et al., 2007; Lukin et al., 2011). Similar situation holds for hyperspectral imaging where noise properties and signal-to-noise ratio (SNR) depend upon sub-band and they vary considerably in different component (sub-band) images (Curran&Dungan, 1989; Uss et al., 2011).

Note that below we mainly focus on considering multichannel images where the general term "multichannel" relates to color, multi- and hyperspectral imaging, dual and multi-polarization radar imaging, multitemporal sensing, where multiple images of the same scene or terrain are obtained. While for one or a few images it is sometimes possible to carry out manual (interactive) image analysis for the determination of noise type and characteristics, it becomes impossible or too labour-consuming to perform such actions for multichannel data, especially if estimation is to be done on-board or under conditions of fast

acquisition and processing of images. Then, one has to apply blind or automatic estimation (Vozel et al., 2009).

It is often assumed that the noise is i.i.d. and either pure additive or pure multiplicative (Petrovic et al., 2007; Ramponi&D'Alvise, 1999). Then, the task of estimating its parameters (variance or relative variance) simplifies since there exist quite many methods able to provide enough accurate estimation (Lukin et al., 2008 and references therein). However, the aforementioned assumption does not hold for a wide range of images formed by modern sensors. Recently, it has been clearly demonstrated (Liu et al., 2008; Foi et al., 2007, Lukin et al., 2011) that even for RGB colour images (for which the model of i.i.d. pure additive noise is still the most popular (Plataniotis&Venetsanopoulos, 2000), this model is not adequate enough. Similarly, for sub-band images of hyperspectral data, the presence of signal-dependent component in addition to additive component has been proved (Uss et al., 2011; Aiazzi et al., 2006) where the signal-dependent component occurs to be dominant for the last generation hyperspectral sensors.

This means that it is necessary to estimate the characteristics of mixed or signal-dependent noise. The corresponding methods are mainly based on forming scatter-plots of local estimates of noise variance (in scanning windows or blocks of a rather small size) on local mean and carrying out robust regression (polynomial curve fitting) into these scatter-plots. Among the methods that belong to this group, it is worth mentioning techniques described in (Liu et al., 2008; Foi et al., 2007; Aiazzi et. al., 2006; Abramov et. al., 2010). They are similar in basic steps as scatter-plot forming and the use of robust fitting curves, but differ in details. A question of how a curve is to be fit in the best or appropriate manner is not discussed in detail. Thus, the main goals of this paper are to consider different approaches to robust regression, to compare their performance, to discuss possible limitations and restrictions, and to give some practical recommendations. Although the problem of robust regression has been studied for several applications (DuMouchel&O'Brien, 1989), to our best knowledge, its use in robust estimation of signal-dependent noise characteristics has not been analyzed thoroughly. Also note that our intention is to attract attention to the problem statement with providing some initial practical solutions rather than developing deep theory.

2. Origins and properties of signal-dependent and mixed noise

By signal-dependent we mean here such a noise that its statistical characteristics (variance, probability density function (PDF)) depend upon information signal (image) in one or another manner. There are quite many known types of signal-dependent noise. Poisson noise (Foi et al., 2007) is the case for which noise variance is equal to the true value of image pixel and noise PDF shape also changes being almost Gaussian for large true values but considerably differing from Gaussian for rather small ones. Another example is film-grain noise commonly assumed locally Gaussian but with variance increasing with image true value (local mean in homogeneous image regions) (Öktem&Egiazarian, 1999). Speckle is one more example for which pure multiplicative model is widely exploited where noise is not Gaussian and its variance quickly increases for larger true values (proportionally to squared local mean in homogeneous image regions) (Oliver&Quegan, 2004; Touzi, 2002; Ramponi&D'Alvise, 1999). Thus, typically a dependence of signal dependent noise variance on true value $\sigma_{sd}^2 = f(I^{tr})$ is monotonically increasing (I^{tr} is true value). The examples

given above show that such properties of noise can originate from different sources, in the first order, from a method of image pixel value obtaining (photon counting, coherent processing of registered signals) or properties of material (carrier) used for data registration.

Mixed noise model (Astola&Kuosmanen, 1997) holds if there are several different sources. The most known example is, probably, mixed additive and impulse noise where the latter component can originate from coding/decoding errors at image transmission via communication channels. Another example is noise in modern sensors where there are such sources as dark noise, thermal noise, photon-counting noise (Kerekes&Baum, 2003). For some sources, noise is signal independent (as dark noise) and for other sources it is signal dependent (as photon-counting noise). However, in aggregate, under assumption of independent noise sources one gets mixed noise for which noise variance occurs to be dependent on image true value as, e.g., $\sigma_{sd}^2 = \sigma_{dc}^2 + kI^{tr}$ where σ_{dc}^2 is variance of dark current noise and k is proportionality factor (Foi et al., 2007). For side look aperture radar images (Lukin et al., 2007), slightly another model of dependence holds $\sigma_{sd}^2 = \sigma_{dc}^2 + k(I^{tr})^2$. Again, dependences $\sigma_{sd}^2 = f(I^{tr})$ are basically monotonously increasing.

However, this is not always true. Different nonlinear transformations of initial data that mainly belong to a wide class of homomorphic and variance stabilizing transforms are often used with special purposes. The intentions of using such transformations are various. For example, logarithmic type transforms can be applied to speckled radar images to convert pure multiplicative to pure additive noise with providing better pre-conditions for applying a wide set of denoising techniques (Oliver&Quegan, 2004; Solbo&Eltoft, 2004). Similarly, Anscombe transform is often used to convert Poisson noise to additive with practically constant variance (Anscombe, 1948). Modified Anscombe-like transforms have been designed to provide variance stabilization for mixed additive and Poisson-like noise in astronomy (Murtagh et al., 1995). Gamma correction is one more example of such nonlinear transformations exploited in digital cameras to improve visual perception of obtained images (Pratt, 2007).

While in the cases of logarithmic and Anscombe transforms it is assumed that original noise type is known and additive noise with constant variance is provided after image homomorphic transforms, the situation with the Anscombe-like transform (Murtagh et al.,1995) and gamma-correction is more complicated.

For carrying out Anscombe-like transform properly, it is needed to know or to estimate parameters or dependence $\sigma_{sd}^2 = \sigma_{dc}^2 + kI^{tr}$. Then, the task reduces to the model described above. In turn, what happens if the standard Anscombe transform is applied to an image corrupted by mixed additive and Poisson noise is considered in the paper (Lukin et al., 2009b). It is demonstrated that dependence of noise variance on local mean becomes monotonously decreasing.

Gamma correction (especially in combination with clipping effects (Foi et al., 2007) can lead to even more specific behaviour of dependence $\sigma_{sd}^2 = f(I^{tr})$. It occurs that in the area of small true values σ_{sd}^2 increases, then, for larger I^{tr}, there is an area of almost constant values of σ_{sd}^2 and finally, for I^{tr} close to an upper limit of dynamic range of image representation, σ_{sd}^2 starts to decrease (Liu et al., 2008; Lim, n.d.; Lukin et al., 2011).

Thus, the dependence $\sigma_{sd}^2 = f(I^{tr})$ is most often monotonously increasing but in special cases it can be also monotonously decreasing or having maximum. This means that there

should be some model of $\sigma_{sd}^2 = f(I^{tr})$ to be fit in a scatter-plot with further estimation of its parameters. Polynomial models with a limited order seem to be a good choice although the use of other quite simple functions is also appropriate. If available, a priori information for model choosing is to be used. Another aspect is availability of methods and algorithms for robust fit of the corresponding curves.

One more question is introducing some restrictions. Since we deal with estimation of noise variance, it should be non-negative by definition. This means that a fitted function $\hat{f}(I^{tr})$ is to be non-negative for entire range of possible values of I^{tr}. Whilst for some functions as, e.g., exponents of different type the non-negativity condition is satisfied automatically, this is not true for polynomial fitting. Therefore, fitting with restrictions is required in practice and this additionally complicates the task. Although selection of regression function is important, below we do not concentrate on it and consider quite simple cases of polynomial fit. It is also worth mentioning here that there exist methods for blind identification of noise/degradation type (Vozel et al., 2006). Recall that these techniques allow identifying additive, multiplicative, impulsive noise, blur and all their possible combinations. However, for the general case of signal-dependent noise considered in this chapter this approach is often useless since noise statistics, as discussed above, can be specific. In particular, these methods are unable to identify Poisson and mixed Poisson and additive noise cases.

3. Robust regression approaches

Several times we have used above the terms "robust regression" and "robust fit" without explaining what is meant by "robust" and why conventional methods of curve fitting into scatter-plots (data) cannot be used. Robustness is treated here in two senses according to Huber (Huber, 1981). First, regression is to be robust with respect to outliers in data. The reasons why outliers appear will be explained in the next subsection. Second, by robustness we also mean the requirement to a method to provide reasonably accurate blind estimation of signal-dependent noise parameters for a wide set of images subject to analysis and a wide range of possible variation of noise statistics.

3.1 Properties of local estimates and ways of scatter-plot pre-processing

Any method for blind estimation of mixed noise parameters starts from obtaining local estimates of noise variance. For this purpose, square shape local windows (blocks) are commonly used. Blocks can be fully overlapping, partially or non-overlapping. In the latter case they are shifted with respect to each other by N pixels in horizontal and/or vertical directions where N denotes block side size. Accuracy of noise parameter estimation provided in the case of non-overlapping blocks is slightly worse (by about two times in the sense of estimated parameter variance), but processing is faster. The method parameter N also influences accuracy. Recommendations concerning its selection are given in (Lukin et al., 2008a) and will be briefly discussed below.

Suppose one has a set of blocks tessellating an analyzed image. Then, local estimates of noise variance are to be obtained for each block or for blocks selected for analysis. Consider first the case of estimate obtaining for all blocks. There are several ways to obtain local estimates. The most known is to calculate

$$\hat{\sigma}_l^2 = (1/(N^2-1)) \sum_{i,j \in G_l} (I_{ij} - \bar{I}_l)^2 , \quad (1)$$

where $\bar{I}_l = (1/N)\sum_{i,j\in G_l} I_{ij}$ denotes the local mean for the l-th block, G_l is the area occupied by the l-th block, I_{ij} is the ij-th pixel of an analyzed image, l=1,…,N_{bl}, N_{bl} is the number of blocks that depends upon image and block size and a way of image tessellating by blocks. Here and below we consider one-channel images assuming that similar operations are carried out for each component image of multichannel data processed sequentially or in parallel).

There are also other algorithms for obtaining local variance estimates. In particular, robust estimators of data scale can be used for this purpose as, for example (Lukin et al., 2005)

$$\hat{\sigma}_l^2 = (1.483(\underset{ij\in G_l}{med}\left|I_{ij} - \underset{ij\in G_l}{med}(I_{ij})\right|))^2, \quad (2)$$

where *med(X)* denotes median value for data sample *X*. However, to apply the local estimate (2) one has to be sure that noise is Gaussian. Otherwise, biased estimates are obtained even in homogeneous image regions. Generally speaking, robust scale estimators (Crnojevic&Petrovic, 2010) can be used for local variance estimation if PDF of signal dependent noise is a priori known and it does not depend on local mean. However, this rarely happens in practice. For example, if signal-dependent noise is Poissonian or Poissonian noise is one component of mixed noise, then noise PDF changes with local mean (in homogeneous image regions).

Other ways to estimate local variance in blocks are possible as well. For example, estimators operating in orthogonal transform domain as, e.g., discrete cosine transform (DCT) can be used (Lukin et al., 2010b). However, these estimators produce biased estimates of noise variance even in image homogeneous regions if noise is spatially correlated (Lukin et al., 2008a, 2010b). Note that we intend on considering both the cases of i.i.d. and spatially correlated noise assuming that no or a limited a priori information is available on noise spatial correlation properties. In practice, noise is often spatially correlated, the reasons for this phenomenon are discussed in (Ponomarenko et al., 2011). To avoid problems dealing with possible biasedness of local estimates, below we focus on considering the estimation algorithm (1) as the basis of scatter-plot forming. Besides, we basically follow recommendations on block size setting given in (Lukin et al., 2008a). According to them, $N\geq 5$ for i.i.d. noise and $N\geq 7$ for spatially correlated noise. However, it is not worth using N>9 in both cases. Thus, N=7, 8, and 9 are good practical choices if a priori information on noise spatial correlation characteristics is not available or is limited.

Several times above it has been mentioned "in homogeneous image regions". This is because just for blocks that belong to homogeneous image regions it is possible to obtain the so-called normal estimates of noise local variance. By normal we mean that such estimates are quite close to the corresponding true value keeping in mind that closeness is determined by block size, PDF of noise and its spatial correlation properties (Lukin et al., 2006, 2008a). Closeness can be characterized by variance of local variance estimates $\sigma^2_{\text{var}} = g(I^{tr})$. For a given I^{tr}, variance σ^2_{var} of normal estimates is directly proportional to $f(I^{tr})$ and it decreases if N increases. For spatially correlated noise, σ^2_{var} is larger than variance of local variance estimates for i.i.d. noise for the same I^{tr}, PDF of noise and N.

Let us demonstrate some of aforementioned properties of local estimates for a very simple test image corrupted by mixed noise (Lukin et al., 2009b). Consider the following case. The test image has the size 512x512 pixels and is composed of 16 horizontal strips each of width

32 pixels. For each strip, I^{tr} is constant and is equal to 20 (for the uppermost strip), 30, 40,…,170, i.e. for an n-th strip its mean is equal to 10+10n, n=1,…,16. The dependence of mixed noise variance is $\sigma_{sd}^2 = \sigma_{add}^2 + I^{tr}$, i.e. additive and Poison noise components are simulated where the latter component is dominant since additive noise variance σ_{add}^2 is set equal to 10 (then $\sigma_{add}^2 < I^{tr}$ for all strips). The noisy image is presented in Fig. 1,a.
The obtained scatter-plot is represented in Fig. 1,b (points in scatter-plot have coordinates $\hat{\sigma}_l^2$ for vertical axis and $\overline{I}_l$ for horizontal axis, N=7, non-overlapping blocks).

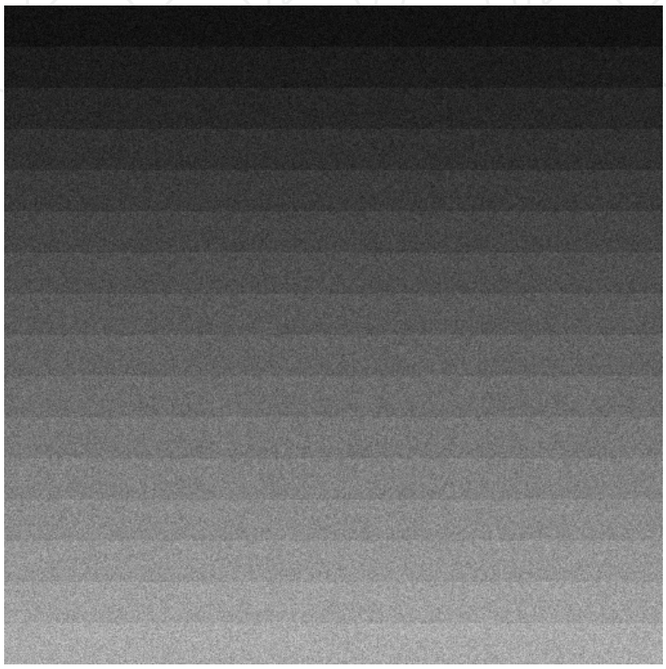


Fig. 1.a Noisy test image

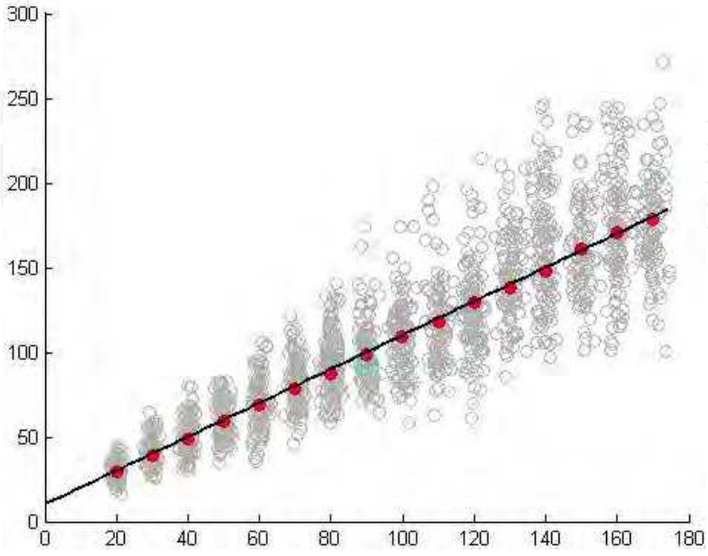


Fig. 1.b Scatter-plot of local estimates for the noisy test image in Fig. 1,a

As it is seen, the obtained points mostly concentrate along the line $\sigma_{sd}^2 = \sigma_{add}^2 + I^{tr}$ also presented at the scatter-plot in Fig. 1,b for convenience of analysis. These points, in fact, form sixteen clusters with the center coordinates approximately equal to (10+*n*, 20+*n*), *n*=1,…,16. These clusters are not of equal size. The clusters that correspond to larger *n* have larger size in both directions (especially, in vertical one) demonstrating that accuracy of estimates of local variance in blocks is worse.

The presented scatter-plot also shows that there are abnormal local estimates (outliers) of local variance placed rather far from any (including the nearest) cluster. These abnormal estimates are observed for blocks that do not fully belong to a given strip, i.e. if a given block falls into edge between strips. In a more general case, abnormal local estimates take place if a block is positioned in heterogeneous image region where by heterogeneity we mean edges, details, and textures.

Below we skip a more detailed analysis of statistics of abnormal local estimates. Only the following is worth stating. First, a percentage of abnormal local estimates basically depends upon structure (complexity) of an analyzed image and usually it is sufficiently larger than for the simple test image in Fig. 1,a (Lukin et al., 2010a). Second, according to experience of blind estimation of additive noise variance or multiplicative noise relative variance (Lukin et al., 2010a), the presence of such abnormal local estimates in any case influence final estimation even if quite robust procedures are applied to joint processing of the set of estimates. Usually, for more complex images and smaller variance of noise the provided accuracy is worse. Third, abnormal estimates give no information on noise properties and one has to rely on normal estimates. However, without special analysis (see (Lukin et al., 2010a) for details) it is difficult to predict in advance for what positions of blocks the normal or abnormal estimates will be obtained.

Intuitively, the presence of outliers should lead to worse accuracy of mixed noise parameter estimation. There are, at least, three ways to diminish the influence of abnormal estimates:

1. to apply aforementioned robust regression in curve fitting to scatter-plots;
2. to determine cluster centres and to carry out curve fitting using only them;
3. to reject introducing into a scatter-plot the estimates for blocks if they are predicted to be abnormal.

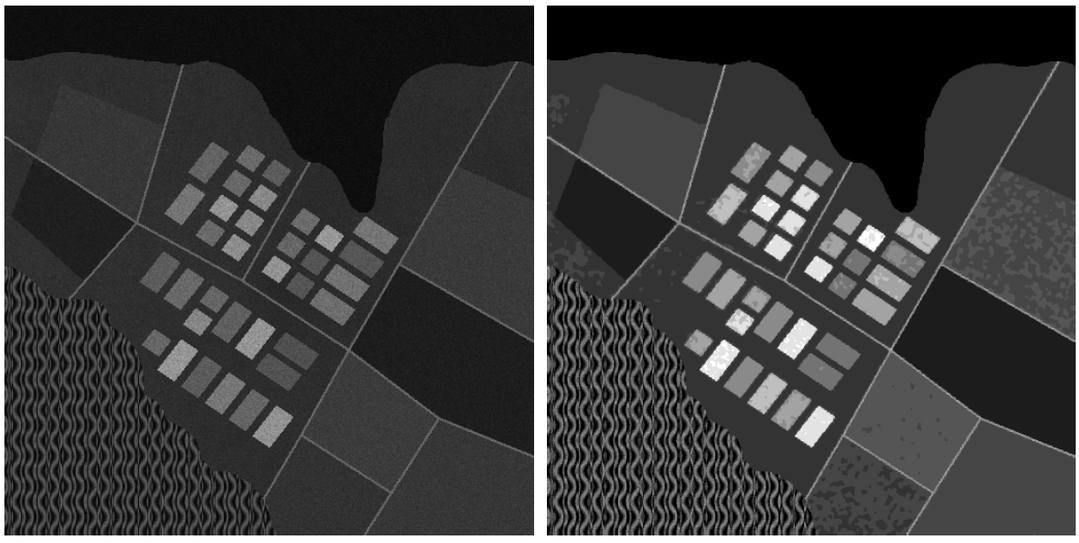
Below we concentrate on considering two former ways. For the way 2, it is possible to apply image pre-segmentation (Klaine et al., 2005). Although there are many methods for image segmentation (Yu-jin Zhang, January 2006), here one needs a method that does not exploit any a priori information on noise type and characteristics. In particular, one can use the method (Klaine et al., 2005) that allows estimating the cluster number and centers. Let us denote them as N_{cl} and $C_{cl\,m}, m = 1,...,N_{cl}$, respectively. Assume that a *k*-th block is referred to an *m*-th cluster if the corresponding central pixel of the block is referred to *m*-th cluster (level) in the segmented image. Thus, we obtain subsets of blocks for each cluster.

For each cluster, a robust method should be applied to determine its center. This method is applied to both a subset of local mean estimates and a subset of local variance estimates. Details can be found in (Lukin et al., 2008b). Thus, one gets cluster center coordinates $\hat{\sigma}_{cl\,m}^2, m = 1,...,N_{cl}$ for vertical axis and $\hat{I}_{cl\,m}, m = 1,...,N_{cl}$ for horizontal axis, respectively. $(\hat{\sigma}_{cl\,m}^2; \hat{I}_{cl\,m}^2), m = 1,...,N_{cl}$. An advantage of this approach is that after estimation of cluster centers the influence of abnormal estimates is reduced radically. The obtained estimates of

cluster centers are shown by red in Fig. 1,b and, as it is seen, they are placed close to the true curve. Then, it is possible to fit a curve (straight line) for getting the estimates $\hat{\sigma}_a^2$ and $\hat{\sigma}_\mu^2$.
Let us give an example. It is taken from the paper (Abramov et al., 2010). The noisy test image RSA is presented in Fig. 2,a. It has been corrupted by mixed additive and multiplicative Gaussian noise mimicking side-look aperture radar images:

$$\sigma_{sd}^2 = \sigma_{add}^2 + \sigma_\mu^2 (I^{tr})^2 = \sigma_{add}^2 + kX,\ X = (I^{tr})^2, \tag{3}$$

where σ_μ^2 denotes relative variance of multiplicative noise. The simulated values are $\sigma_{add}^2 = 9, \sigma_\mu^2 = 0.005$. While forming a scatter-plot, $(I^{tr})^2$ has been replaced by X for horizontal axis in order to have an opportunity to fit a first-order polynomial (such "tricks" are possible if one is confident that dependence is as (3)). The pre-segmented image is represented in Fig. 2,b.



a b

Fig. 2. Noise test image RSA (a) and the result of its pre-segmentation (b)

The obtained scatter-plot is represented in Fig. 3,a. The cluster centers are marked by red squares. For larger X, clusters are of larger size as in the previous case of signal-dependent noise with variance monotonously increasing with I^{tr}. Abnormal estimates are observed as well especially for the clusters with relatively small X. Green line shows the true dependence $\sigma_{sd}^2 = 9 + 0.005X$ whilst the red one corresponds to the curve fitted by the method (Lukin et al., 2008b) which is LMS fit using cluster centers. As seen, the curves are quite close but anyway they differ from each other.

3.2 Scatter-plot and curve fitting peculiarities for real life data

Above we have analyzed test images which are quite simple (they contain rather homogeneous image regions of rather large size). In this subsection, we consider some examples for real life data. The first example is taken from TerraSAR-X data, the image is presented in Fig. 4,a.

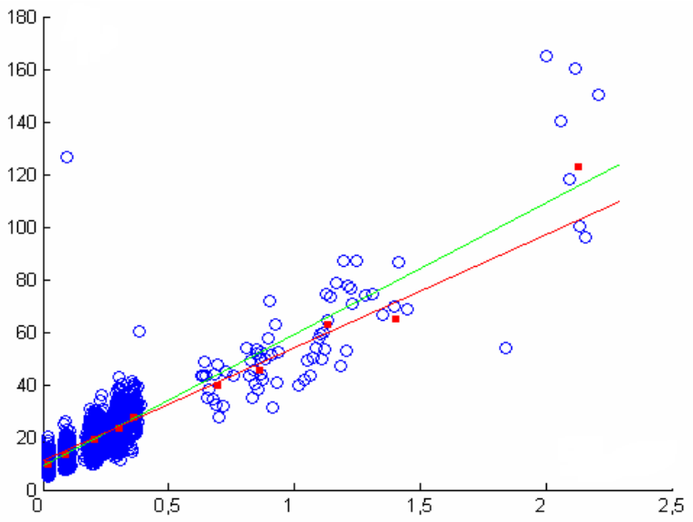
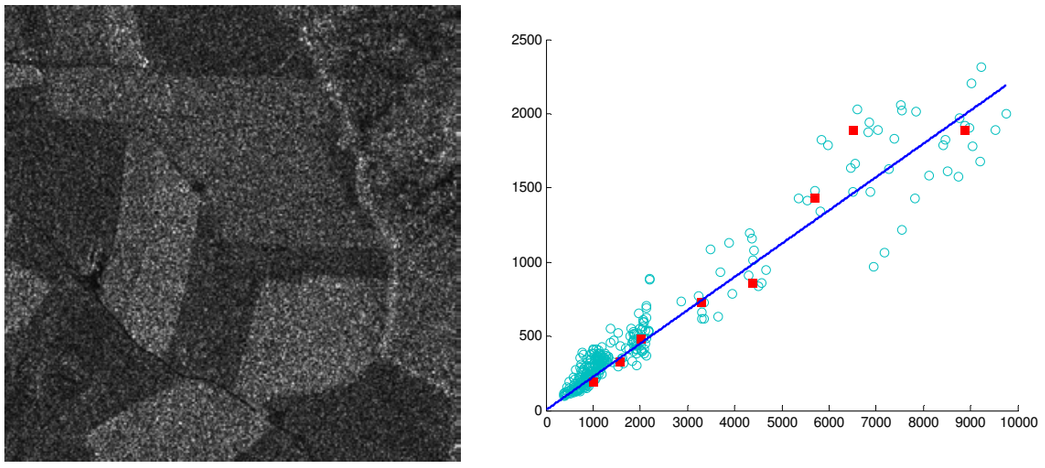


Fig. 3. a Scatter-plot of local estimates for the noisy test image in Fig. 2,a



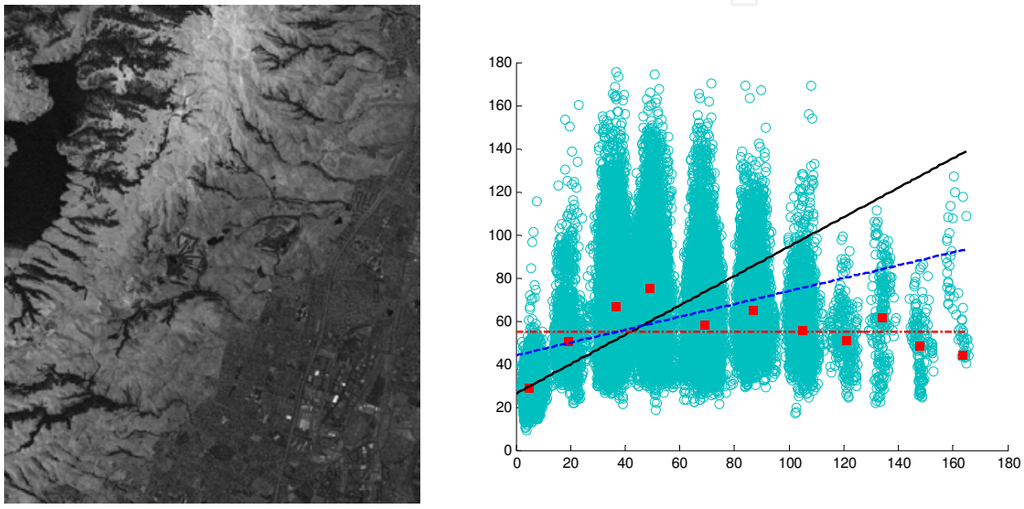
(a) (b)

Fig. 4. Real life SAR image (a) and the scatter-plot of local variance estimates after pre-segmentation (b)

This is a one-look amplitude SAR image of agricultural region in Germany (Rosenheim, http://www.infoterra.de) with fully developed speckle for which multiplicative noise is dominant and its PDF is close to Rayleigh. Since we neglect the influence of additive noise, the line that passes through the coordinate origin has been fitted. The estimated $\hat{\sigma}_{\mu}^{2}$ is equal to 0.22, i.e. the estimate is quite close to the value 0.273 typical for speckle with Rayleigh PDF. This example shows that, if there is reliable a priori information on noise properties, one possible practical restriction of regression when additive noise variance is supposed equal to zero.

Even more interesting results are given in Fig. 5. Fig. 5,a presents the 168-th sub-band image of hyperspectral AVIRIS data Moffett Field 1 (http://aviris.jpl.nasa.gov). Recall that it has been supposed that for hyperspectral data signal-dependent noise is characterized by the

dependence $\sigma_{sd}^2 = \sigma_{dc}^2 + kI^{tr}$. Thus, the scatter-plot of local variance estimates has been obtained (see Fig. 5,b) and regression has been carried out using three techniques. The first is Robust fit (RLMS) available in Matlab (black line). The second is standard LMS fit carried out for cluster centers (red line), and the third one is weighted LMS (WLMS) (Abramov et al., 2010) also applied to cluster centers. The estimates of additive noise variance are equal to 26.3, 54.9, and 43.9, respectively. The estimates of the parameter k are equal to 0.68, -0.00024, and 0.3, respectively. Whilst for additive noise variance estimates differ from each other but not too much, the estimates of k are very different. Moreover, for the LMS regression the obtained estimate is negative. If we suppose that the dependence is monotonically increasing, the estimate should be positive and this can be imposed as restriction to curve fitting algorithm.



(a) (b)

Fig. 5. Real life sub-band image of hyperspectral data (a) and the scatter-plot of local variance estimates with fitted lines (b)

It is interesting that for the same 168-th sub-band but for another hyperspectral data set (Lunar Lake) the estimates of additive noise variance are 27.8, 28.5 and 24.1, respectively (for the RLMS, LMS, and WLMS methods). The estimates of k are 0.07, 0.08, and 0.1. Thus, in this case, the estimates almost coincide for all three methods, and this indirectly indicates that these estimates are accurate enough. Meanwhile, these results do not give answer to a question what method is more accurate. The presented examples show only that different regression techniques can produce either quite similar or very dissimilar estimates of mixed noise parameters. Thus, numerical simulation for a set of test images of different complexity and a set of mixed noise parameters is needed to get imagination on what regression methods are better and to provide practical recommendations.

4. Numerical simulation results

To carry out numerical simulations, we have to select methods for analysis and comparisons, quantitative criteria, test images and sets of mixed noise parameters. Let us at the very beginning give brief description of regression techniques.

4.1 Considered regression techniques

In our study, we have considered the following regression techniques. The first technique is the RobustFit method realized in Matlab that is applied to scatter-plot of local estimates without their pre-processing. The second is the standard LMS fit applied to cluster centers. For this technique, robustness to outliers is provided due to robust determination of cluster centers. The third technique is weighted LMS (Abramov et al., 2010) applied to cluster centers. Its basic idea is to assign weights inversely proportional to the number of points in each cluster keeping in mind that accuracy of cluster center determination increases if the number of such points becomes larger. The fourth technique is Ransac (Fischler&Rolles, 1981) that employs lines fitting for arbitrary pairs of cluster centers, forming a confidence interval and rejecting cluster centers that occur to be out of the confidence interval. In this way, the cluster centers assumed to be determined with the worst accuracy are removed. The fifth technique is the so-called double weighted LMS (DWLMS) that has two stages. At the first stage, WLMS is applied and then, at the second stage, the cluster centers that are far from the fitted line are taken into account with smaller weights to additionally improve accuracy.

As it is seen, four the latter techniques operate with the cluster centers and this means that accuracy of cluster center determination affects the final accuracy of these techniques. To get initial imagination on the influence of cluster center estimation accuracy, let us carry out simple analysis of estimation accuracy of mixed noise parameters for the case of curve fitting for two points. It is worth mentioning here that such line can be fitted uniquely and such fitting is the main operation at the first stage of Ransac technique.

4.2 Accuracy of line fitting for two points

Suppose that we have two points (e.g., cluster centers) that have coordinates $x_1+\Delta x_1$, $y_1+\Delta y_1$ and $x_2+\Delta x_2$, $y_2+\Delta y_2$ where x_1, y_1 and x_2, y_2 denote the true values whilst Δx_1, Δy_1 and Δx_2, Δy_2 are the errors of point (cluster center) determination. Assume also that the following conditions are valid: $y_2+\Delta y_2>0, y_1+\Delta y_1>0$ and $x_2+\Delta x_2>x_1+\Delta x_1>0$. We suppose also that $\frac{\Delta y_2}{y_2}\le 1$, $\frac{\Delta y_1}{y_1}\le 1$, $\frac{\Delta x_2}{x_2}\le 1$, $\frac{\Delta x_1}{x_1}\le 1$, i.e. the errors are comparatively small (validity of these assumptions in practice will be discussed later). Then, fitted line parameters a and b are to be determined from the following simple equation system

$$\begin{cases} y_1+\Delta y_1 = a+b(x_1+\Delta x_1) \\ y_2+\Delta y_2 = a+b(x_2+\Delta x_2) \end{cases}, \tag{4}$$

Then one has

$$\hat{b}=\frac{y_2+\Delta y_2-(y_1+\Delta y_1)}{x_2+\Delta x_2-(x_1+\Delta x_1)}=\frac{y_2-y_1+(\Delta y_2-\Delta y_1)}{x_2-x_1+(\Delta x_2-\Delta x_1)}=\frac{y_2-y_1+(\Delta y_2-\Delta y_1)}{(x_2-x_1)\left(1+\frac{\Delta x_2-\Delta x_1}{x_2-x_1}\right)}, \tag{5}$$

and, under introduced assumptions on relatively small errors

$$\hat{b}\approx\frac{y_2-y_1+(\Delta y_2-\Delta y_1)}{x_2-x_1}\left(1-\frac{\Delta x_2-\Delta x_1}{x_2-x_1}\right)=\left(\frac{y_2-y_1}{x_2-x_1}+\frac{\Delta y_2-\Delta y_1}{x_2-x_1}\right)\left(1-\frac{\Delta x_2-\Delta x_1}{x_2-x_1}\right) \tag{6}$$

Let us introduce notations $\Delta x = \Delta x_2 - \Delta x_1$, $\Delta y = \Delta y_2 - \Delta y_1$. In this case

$$\begin{aligned}\hat{b} &\approx \left(b_{true} + \frac{\Delta y}{x_2 - x_1} \right)\left(1 - \frac{\Delta x}{x_2 - x_1} \right) = b_{true} - b_{true}\frac{\Delta x}{x_2 - x_1} + \frac{\Delta y}{x_2 - x_1} - \frac{\Delta x \Delta y}{x_2 - x_1}, \\ &\approx b_{true} + \frac{\Delta y - b_{true}\Delta x}{x_2 - x_1}\end{aligned} \tag{7}$$

where b_{true} is the true value of parameter *b*. Hence

$$\Delta b = \frac{\Delta y - b_{true}\Delta x}{x_2 - x_1}. \tag{8}$$

Then, assuming that $\sigma^2_{\Delta y_1}$, $\sigma^2_{\Delta y_2}$, $\sigma^2_{\Delta x_1}$, and $\sigma^2_{\Delta x_2}$ are variances of the corresponding cluster center coordinates, it is easy to obtain variance of parameter b estimate as

$$\sigma^2_{\Delta b} = \frac{\sigma^2_{\Delta y_1} + \sigma^2_{\Delta y_2} + b^2_{true}(\sigma^2_{\Delta x_1} + \sigma^2_{\Delta x_2})}{(x_2 - x_1)^2}. \tag{9}$$

This expression shows the following. Since b_{true} is usually smaller than unity and coordinates *x* are estimated with rather high accuracy, the main contribution to errors of b estimation results from Δy_1 and Δy_2. Moreover, if these errors do not have zero mean (below and in (Lukin et al., 2008a) it is demonstrated that this happens often), biased estimation of b takes place. An important conclusion is also that $\sigma^2_{\Delta b}$ is inversely proportional to $(x_2 - x_1)^2$. This is intuitively clear that more distant points "fix" a fitted line better, but this property is in no way exploited in robust regression. One more conclusion is that for smaller b_{true} variance $\sigma^2_{\Delta b}$ also decreases. Meanwhile, the ratio $\frac{\sigma^2_{\Delta b}}{b^2_{true}}$ characterizing relative error increases.

Similarly, it is possible to estimate parameter a as

$$\hat{a} = \frac{1}{2}\left(y_1 + \Delta y_1 + y_2 + \Delta y_2 - b(x_1 + \Delta x_1 + x_2 + \Delta x_2) \right). \tag{10}$$

Then one gets

$$\begin{aligned}\hat{a} &= \frac{1}{2}\left(y_1 + y_2 - b_{true}(x_1 + x_2) \right) + \frac{1}{2}\left(\Delta y_1 + \Delta y_2 - b_{true}(\Delta x_1 + \Delta x_2) - \Delta b(x_1 + \Delta x_1 + x_2 + \Delta x_2) \right) = \\ &= a_{true} + \Delta a\end{aligned} \tag{11}$$

where a_{true} is the true value of parameter a and

$$\Delta a = \frac{1}{2}\left(\Delta y_1 + \Delta y_2 - \left(b_{true}(\Delta x_1 + \Delta x_2) + \frac{\Delta y - b_{true}\Delta x}{x_2 - x_1}(x_1 + \Delta x_1 + x_2 + \Delta x_2) \right) \right). \tag{12}$$

Then variance of estimation is

$$\sigma^2_{\Delta a} = \frac{1}{4}\left(\sigma^2_{\Delta y_1} + \sigma^2_{\Delta y_2} + (\sigma^2_{\Delta y_1} + \sigma^2_{\Delta y_2})\frac{(x_1 + x_2)^2}{(x_2 - x_1)^2} + \frac{4b^2_{true}(\sigma^2_{\Delta x_1} + \sigma^2_{\Delta x_2})}{(x_2 - x_1)^2} \right). \tag{13}$$

Analysis of this expression shows the following. As can be easily predicted, $\sigma^2_{\Delta a}$ increases if variances $\sigma^2_{\Delta y_1}$, $\sigma^2_{\Delta y_2}$, $\sigma^2_{\Delta x_1}$, and $\sigma^2_{\Delta x_2}$ become larger. Thus, it is desirable to use cluster centers determined with the best accuracy. Besides, again estimation variance decreases if points are selected as distantly as possible (with larger $(x_2 - x_1)^2$ and smaller $(x_2 + x_1)^2$, i.e. if x_1 is as close to zero as possible). Finally, $\sigma^2_{\Delta a}$ is smaller if absolute values of b_{true} approach to zero (this property shows that the estimates of parameters a and b are mutually dependent). It is also easy to show that nonzero mean values of Δx_1 and Δx_2 and, especially, of Δy_1 and Δy_2 result in biased estimation.
The performed analysis shows why it is desirable to carry out weighted robust fit into pre-processed scatter-plot data. However, there are quite many ways to do this. Comparative analysis results will be presented in the next section.

4.3 Test images and accuracy criteria

The study has been performed for three test images, namely, the RSA image given in Fig. 2 and the standard test images Peppers and Goldhill, all of size 512x512 pixels. The test images Peppers and, especially, Goldhill are more complex than the RSA image. Thus, their joint and comparative analysis allows analyzing the influence of image complexity on accuracy of mixed noise parameter estimation.
Simulations, without loosing generality, have been carried out for the model for mixed additive and multiplicative noise $I_{ij} = I^{tr}_{ij} \cdot \mu_{ij} + n_{ij}$ where I^{tr}_{ij} denotes noise-free image, μ_{ij} defines multiplicative noise component obeying Gaussian distribution with unity mean and relative variance σ^2_{μ}, and n_{ij} describes additive noise component with zero mean Gaussian distribution with variance σ^2_a.
As a quantitative criteria, the estimation bias

$$\Delta_x = \left| \left\langle \hat{\sigma}^2_x \right\rangle - \sigma^2_x \right|, \tag{14}$$

variance

$$\theta^2_x = \left\langle \left(\hat{\sigma}^2_x - \left\langle \hat{\sigma}^2_x \right\rangle \right)^2 \right\rangle, \tag{15}$$

and aggregate error

$$\varepsilon_x = \Delta^2_x + \theta^2_x, \tag{16}$$

have been used. Sub-index x denotes belonging of the corresponding parameter for additive (a) or multiplicative (μ) noise. Notation $\langle \bullet \rangle$ means averaging by realizations. To provide statistically stable results the number of realizations was 100.

4.4 Simulation result analysis

Two sets of mixed noise parameters have been used in simulations. The first that can be called non-intensive noise is $\sigma_a^2 = 10$, $\sigma_\mu^2 = 0.005$. The second case is $\sigma_a^2 = 100$, $\sigma_\mu^2 = 0.05$ where for both sets multiplicative noise becomes dominant for image true values over 45.
The obtained results are given in Table 1. The optimal parameter for the RSC technique is given after /.

Let us start from considering the simplest test image RSA. As it is seen, the bias Δ_a is very small (the absolute value of the ratios Δ_a / σ_a^2 are less than 0.009 for all five considered regression techniques. Meanwhile, the values θ_a^2 considerably differ for the analyzed methods and just θ_a^2 mainly contributes to ε_a (except the results for DWLMS). The values

Image	Para-meters	Method	Δ_a	θ_a^2	ε_a	$\Delta_\mu \times 10^{-4}$	$\theta_\mu^2 \times 10^{-8}$	$\varepsilon_\mu \times 10^{-8}$
RSA	$\sigma_a^2 = 10$ $\sigma_\mu^2 = 0.005$	RLMS	0.0181	0.0080	0.0083	-1.1	0.48	1.7
		LMS	0.038	0.027	0.028	-1.5	1.0	3.0
		WLMS	0.094	0.0312	0.040	-1.8	1.0	5.0
		RSC / 7	0.0199	0.031	0.0315	-1.4	1.0	3.0
		DWLMS	-0.086	0.0069	0.014	-0.93	0.1	1.0
	$\sigma_a^2 = 100$ $\sigma_\mu^2 = 0.05$	RLMS	1.02	0.84	1.89	-23	59	590
		LMS	-1.28	2.48	4.12	-14	96	280
		WLMS	-0.93	3.64	4.50	-16	140	390
		RSC / 9	-1.27	3.84	5.45	-14	150	330
		DWLMS	-1.63	0.71	3.35	-13	38	20
Peppers	$\sigma_a^2 = 10$ $\sigma_\mu^2 = 0.005$	RLMS	15.5	2.76	243.5	-4.5	3.5	24
		LMS	19.16	1.25	368.43	-3.1	1.0	11
		WLMS	15.4	2.23	239.36	-3.5	2.0	14
		RSC / 9	20.8	39.1	469	-4.2	5.0	23
		DWLMS	12.04	0.83	145.87	-1.8	1.0	4.0
	$\sigma_a^2 = 100$ $\sigma_\mu^2 = 0.05$	RLMS	31.2	78.4	1052	-24	70	630
		LMS	18.6	83.1	429.15	-12	79	230
		WLMS	14.53	58.13	269.33	-11	49	170
		RSC / 7	13.6	46.9	232	12	40	180
		DWLMS	8.42	5.86	76.77	-8.2	22	89
Goldhill	$\sigma_a^2 = 10$ $\sigma_\mu^2 = 0.005$	RLMS	32.01	1.9	1027	-5.8	0.56	33
		LMS	27.89	1.55	779.3	-3.7	1.0	14
		WLMS	27.91	1.58	780.54	-4.8	1.0	23
		RSC / 6	28.1	8.5	801	-4.8	1.0	25
		DWLMS	18.96	0.86	360.43	-1.5	1.0	3.0
	$\sigma_a^2 = 100$ $\sigma_\mu^2 = 0.05$	RLMS	62.0	170.7	4015	-2.1	67	520
		LMS	29.74	66.11	950.8	-14	49	210
		WLMS	29.37	79.91	942.61	-13	67	240
		RSC / 8	24.0	94	671	-8.8	93	170
		DWLMS	17.42	9.95	313.46	-8.6	37	110

Table 1. Comparative results analysis for test images

ε_a and the ratios $(\varepsilon_a)^{1/2}/\sigma_a^2$ characterizing aggregate relative accuracy are small for all techniques (smaller than 0.02 that can be considered acceptable in practice).

For the multiplicative component, the bias for all methods is negative and the absolute values of the ratios $\Delta_\mu / \sigma_\mu^2$ are not larger than 0.036. The ratios $(\varepsilon_\mu)^{1/2}/\sigma_\mu^2$ do not exceed 0.048. This does not cause serious problems in practice (Abramov et al., 2004). The main contribution to ε_μ results from estimation bias although contribution of θ_μ^2 is also sufficient.

The same conclusions hold for the case $\sigma_a^2 = 100$, $\sigma_\mu^2 = 0.05$. The difference is that the absolute values of all quantitative criteria are larger. However, the ratios $(\varepsilon_a)^{1/2}/\sigma_a^2$ are still small enough (less than 0.024). The ratios $\Delta_\mu / \sigma_\mu^2$ and $(\varepsilon_\mu)^{1/2}/\sigma_\mu^2$ are of the same order. Thus, the provided accuracy is acceptable for practice (Lukin et al., 2009a).

Consider now the test image Peppers that is slightly more complex than RSA. Let us start with the case $\sigma_a^2 = 10$, $\sigma_\mu^2 = 0.005$. For additive noise component, the bias Δ_a is large and positive for all robust regression techniques. The values of the ratios Δ_a / σ_a^2 vary from 1.2 for DWLMS to about 2.1 for RSC. This shows that additive noise variance is overestimated. Estimation bias contribution to aggregate error ε_a is dominant.

On the contrary, multiplicative noise variance is underestimated (for all methods Δ_μ is negative and its contribution to ε_μ is dominant. The absolute values of the ratios $\Delta_\mu / \sigma_\mu^2$ are not larger than 0.07. The ratios $(\varepsilon_\mu)^{1/2}/\sigma_\mu^2$ do not exceed 0.1. The best results are provided by the method DWLMS.

For the case $\sigma_a^2 = 100$, $\sigma_\mu^2 = 0.05$, the values of bias Δ_a are large enough and positive for all analyzed robust regression techniques. However, the ratios Δ_a / σ_a^2 are smaller than in the previous case. Again, the influence of bias (systematic error) on ε_a is dominant. The absolute values of the ratios $\Delta_\mu / \sigma_\mu^2$ and $(\varepsilon_\mu)^{1/2}/\sigma_\mu^2$ are smaller than in the previous case. The technique DWLMS provides the best accuracy.

Finally, consider the test image Goldhill. If $\sigma_a^2 = 10$, $\sigma_\mu^2 = 0.005$, large positive valued bias is observed for estimates of additive noise variance. Multiplicative noise variance is estimated well enough, although it is slightly underestimated. Sufficient bias takes place for the estimates of additive noise variance if $\sigma_a^2 = 100$, $\sigma_\mu^2 = 0.05$. Multiplicative noise variance is again underestimated. The best accuracy is provided for the method DWLMS whilst the worst accuracy is observed for RMLS in almost all cases.

Summarizing the obtained results, it is possible to conclude the following:

1. the estimates of additive noise variance are usually biased and overestimated;
2. the estimates of multiplicative noise variance are also biased but underestimated;
3. bias contributes to aggregate error more than estimation variance, thus, its reduction is the first order task;
4. estimation accuracy is worse for more complex images; the ratios $(\varepsilon_\mu)^{1/2}/\sigma_\mu^2$ and $(\varepsilon_a)^{1/2}/\sigma_a^2$ are larger for the case of less intensive noise, thus it is more difficult to provide appropriate accuracy just for non-intensive noise situations;
5. the method DWLMS usually provides the best accuracy.

To our opinion, the main drawback of all considered techniques is overestimation (positive bias) of additive noise component. There could be three main reasons for this phenomenon. The first is self-noise in test images. The second is the influence of local image content on local variance estimates. The third is the influence of heavy tail of distributions (abnormal estimates) present in clusters. Therefore, it is worth trying to decrease this bias. However,

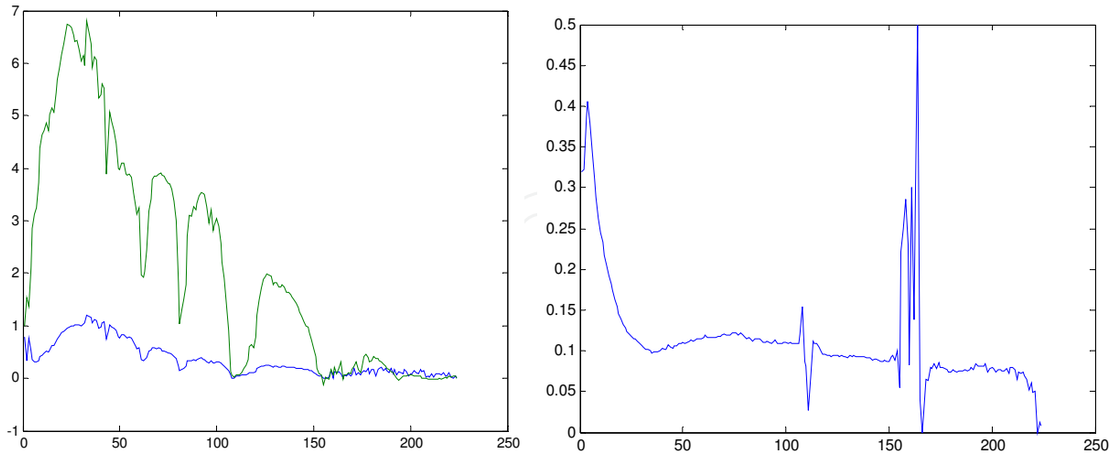
the decision to use the smallest local estimates (Liu et al., 2008) does not seem the best solution since such estimates can be by several times smaller than the true value of mixed noise variance for a given local mean.

5. Real life data testing

Testing of the methods performance has been also carried out for real life AVIRIS images for which the assumed model of mixed noise is described as $\sigma_{sd}^2 = \sigma_{add}^2 + kI^{tr}$. Thus, we have to estimate σ_{add}^2 and k . The scatter-plot example has been earlier given in Fig. 5.

We have applied the Robustfit method component-wise to two AVIRIS images, namely, Lunar Lake and Moffett Field 1. For the Moffett Field image, it has produced quite many (more than 15%) negative values of the estimates $\hat{\sigma}_{add}^2$ and a very wide range of these estimates (from -180000 to 190000). For the Lunar Lake image, no negative valued estimates $\hat{\sigma}_{add}^2$ have been obtained but the largest values have been up to 62000. Clearly, such accuracy is inappropriate since they do not agree with the estimates for other methods.

If pre-segmentation method (Klaine et al., 2005) is applied to each component image with removal of heterogeneous blocks from further consideration and then Robustfit is used, the estimates $\hat{\sigma}_{add}^2$ become in better agreement with the estimates produced by other techniques. At least, the number of negative valued estimates reduces and the limits of their variation become considerably narrower.



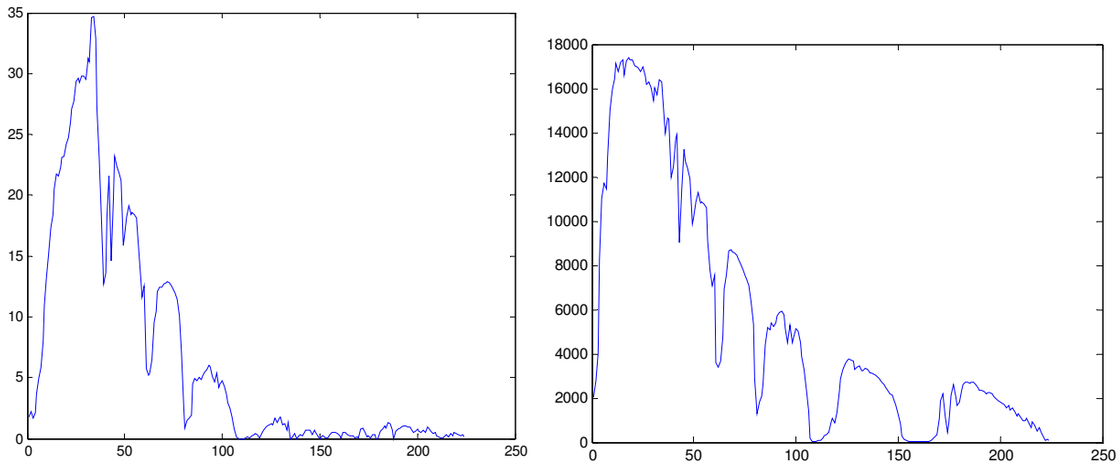
(a) (b)

Fig. 6. Estimates of *k* for Robustfit (green) and DWLMS (blue) techniques (a) and dependence of $I_{\min} / I_{\max}$ on sub-band index (b)

The estimates of k for the Lunar Lake data for the Robustfit method with pre-segmentation by the method (Klaine et al., 2005) are presented in Fig. 6,a (green color). As it is seen, mostly they are larger than unity and only for a small percentage of sub-bands, mostly with indices larger than 160, the estimates of k are smaller than unity. These results are not in agreement with results in other papers dealing with estimation of noise characteristics in AVIRIS hyperspectral images (see Uss et al., 2011, and references therein). Just overestimation of k can lead to underestimation of σ^2_{add} and negative values of its estimated mentioned earlier.

This example shows that it is worth imposing restrictions on non-negativity of both the estimates of σ^2_{add} and k. The DWLMS technique with imposed restrictions has produced quite many zero estimates of σ^2_{add}. The provided estimates of k occur to be considerably smaller (shown by blue color curve in Fig. 6,a) than for the Robustfit. Meanwhile, the shapes of these curves are very similar.

We have also analyzed dynamic range of data in sub-band images with determining the minimal and maximal values $I_{\min}$ and $I_{\max}$ in each sub-band image. The ratios $I_{\min} / I_{\max}$ for all 224 sub-bands of hyperspectral data Lunar Lake are represented in Fig. 6,b. It is seen that for most sub-bands the ratios are not zero. Thus, the histograms of sub-band image values have specific behaviour compared to most optical test images (that usually have quite many values close to zero). Then, it is difficult to expect that there are clusters that have relatively small $\hat{I}_{clm}$ and, according to analysis in subsection 4.2, the estimation accuracy reduces.



(a) (b)

Fig. 7. Estimates of *k* for DWLMS technique (a) and dependence of $I_{\max}$ on sub-band index (b)

It can be observed from Fig. 6,a that the estimates for neighbour sub-bands (close values of sub-band wavelengths) are close to each other. The estimates of k are also in correlation with sub-band image dynamic range. This is seen from analysis of plots in Fig. 7. The first plot (Fig. 7,a) shows dependence of $\hat{k}$ on sub-band index for the AVIRIS data Moffett Field 1 (the estimates have been obtained by the DWLMS method with imposed restrictions). The second plot presents dependence of $I_{\max}$ on sub-band index. It is seen that the curves have quite similar shapes.

6. Conclusions and future work

The study carried out above allows drawing a set of conclusions and giving a list of possible directions of future work.

First, estimation of mixed and signal-dependent noise parameters is only at the beginning of its development. The scatter-plot or cluster-center representations are the basis for other operations (curve regression) applied at next stages in any case. The problem of standard scatter-plot approach is that it usually contains abnormal estimates that influence estimation accuracy for any robust regression technique applied. Thus, it is desirable "to cope" with abnormal local estimates at initial stage of data processing attempting to reject them. However, it is not an easy task if parameters of mixed noise are unknown. In turn, there are also problems for cluster-center based representation. The questions that arise are: a) how to select cluster number? b) how to estimate their center positions with appropriate accuracy? c) how to predict accuracy of such estimation?

Second, the studies with simulated noise for test images have shown that even the local estimates considered normal can be considerably biased. This drawback is especially of value for additive noise component if its variance is not large. The relative bias characterized by Δ_a / σ_a^2 can be quite large. The main reason is the influence of image content. This means that one way to improve estimation accuracy is to design more accurate and robust methods for estimating local variance with diminishing the influence of image information content.

Third, the weighted methods of LMS regression using cluster centers have demonstrated their advantages. However, the potential benefits of these methods seem to be not exploited in full extent. Currently only the number of points in clusters and distances from cluster centers to initially fitted curve are used in weight adapting. It seems expedient to take other properties of clusters into account as well. One such property could be cluster size characterized in a robust manner. The positions of cluster centers and distances between them can be taken into account as well. Analysis of these aspects can be one more direction of future research.

Fourth, a priori information on mixed or signal-dependent noise is of great importance. We have carried out our experiments supposing that a model of mixed noise is a priori known and it is valid. In practice, it can be only known that noise is signal-dependent but a character (properties) of such dependence can be unknown. Then, a question arises what curve to fit? Polynomials of low order seem to be a natural choice at the first glance. However, one should keep in mind that a fitted polynomial might have intervals of negative values which, according to definition of noise local variance, is not acceptable. One simple way out is to replace negative values of a fitted polynomial by zeroes but is this the best way? Thus, the following problems and questions still remain: 1) how to impose restrictions

on fitted curves and their parameters? 2) what models of curves to apply? 3) what order of curves (polynomials) to use?

Fifth, our experiments with simulated noise have been performed only for i.i.d. noise. Spatial correlation of noise leads to several specific outcomes. In particular, statistics of local estimates changes. Certainly, this influences the performance of entire procedure of mixed noise parameters' estimation. This means that the studies for mixed noise with essential spatial correlation are to be carried out in future. Besides, if estimation is performed for hyperspectral data, considerable correlation of noise statistics in neighbour sub-bands is worth using to improve estimation accuracy.

Sixth, the goal of estimating mixed noise parameters is to use the obtained estimates at later stages of image processing. Operations used at later stages can be homomorphic transforms, edge detection, filtering, lossy compression, etc. Since any estimation is not perfect, the estimation errors influence performance of methods and algorithms applied at later stages. Degree of such negative influence is to be investigated and this will allow formulating practical requirements to accuracy of parameter estimation for mixed noise.

We also see other directions of research and studies. Whilst for particular cases of mixed and signal dependent noise there exist variance-stabilizing transforms, the general theory of such transforms is far from completeness. Note that the use of variance-stabilizing transforms simplifies applying many existing image processing methods and algorithms. In this sense, the recent studies (Foi, 2009) show perspectives and directions of future work.

7. References

Abramov S., Lukin V., Ponomarenko N., Egiazarian K., & Pogrebnyak O. (2004). Influence of multiplicative noise variance evaluation accuracy on MM-band SLAR image filtering efficiency. *Proceedings of MSMW 2004*, Vol. 1, pp. 250-252, Kharkov, Ukraine, June 2004

Abramov S., Zabrodina V., Lukin V., Vozel B., Chehdi K., & Astola J. (2010). Improved method for blind estimation of the variance of mixed noise using weighted LMS line fitting algorithm. *Proceedings of ISCAS*, pp. 2642-2645, Paris, France, June 2010

Aiazzi B., Alparone L., Barducci A., Baronti S., Marcoinni P., Pippi I., & Selva M. (2006). Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers. *Annals of Geophysics*, Vol. 49, No. 1, February 2006

Al-Shaykh O.K. & Mersereau R.M. (1998). Lossy Compression of Noisy Images, *IEEE Transactions on Image Processing,* Vol. 7, No 12, (December 1998), pp. 1641-1652.

Anscombe F.J. (1948). The transformation of Poisson, binomial and negative binomial data. *Biometrica*, Vol. 35, pp. 246-254

Astola J. & Kuosmanen P. (1997). *Fundamentals of nonlinear digital filtering*, CRC Press LLC, Boca Raton, USA

Bovik A. (2000). *Handbook on Image and Video Processing.* Academic Press, USA

Crnojevic V. & Petrovic N. (2010). Impulse Noise Filtering Using Robust Pixel-Wise S-estimate of Variance. *EURASIP Journal on Advances in Signal Processing*, Volume 2010, Article ID 830702, doi:10.1155/2010/830702

Curran P.J. & Dungan J.L. (1989). Estimation of signal-to-noise; a new procedure applied to AVIRIS data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 27, pp. 620-628

Donoho D.L. (1995). De-noising by Soft Thresholding. *IEEE Trans. on Information Theory,* Vol. IT-41, No 3, pp. 613-627

DuMouchel W. & O'Brien F. (1989). Integrating a Robust Option into a Multiple Regression Computing Environment in Computing Science and Statistics. *Proceedings of the 21st Symposium on the Interface*, pp. 297-301, American Statistical Association, Alexandria, VA

Elad M. (2010). *Sparse and Redundant Representations. From Theory to Applications in Signal and Image Processing*, Springer Science+Business Media, LLC

Fischler M.A. & Rolles R.C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, Vol. 24, No. 6, pp. 381-395

Foi A., Trimeche M., Katkovnik V., & Egiazarian K. (2007). Practical Poissonian-Gaussian Noise Modeling and Fitting for Single Image Raw Data. *IEEE Transactions on Image Processing*, Vol. 17, No. 10, pp. 1737-1754

Foi A. (2009). Clipped Noisy Images: Heteroskedastic Modeling and Practical Denoising. *Signal Processing*, Vol. 89, No. 12, pp. 2609-2629

http://aviris.jpl.nasa.gov

Huber P. (1981). *Robust statistics*, Wiley, New York

Kerekes J.P. & Baum J.E. (2003). Hyperspectral Imaging System Modeling. *Lincoln Laboratory Journal*, Vol. 14, No. 1, pp. 117-130

Klaine L., Vozel B., & Chehdi K. (2005). Unsupervised Variational Classification Through Image Multi-Thresholding. *Proceedings of the 13th EUSIPCO Conference*, Antalya, Turkey

Lim S.H. (2006). Characterization of Noise in Digital Photographs for Image Processing. *Proceedings of Digital Photography II*, SPIE 6069, DOI: 10.1117/12.655915

Liu C., Szeliski R., Kang S.B., Zitnick C.L., & Freeman W.T. (2008). Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No 2, pp. 299-314

Lukin V., Koivisto P., Ponomarenko N., Abramov S., & Astola J. (2005). Two-stage Methods for Mixed Noise Removal. *CD-ROM Proceedings of EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP)*, Japan, May 2005

Lukin V., Abramov S., Ponomarenko N., Vozel B., & Chehdi K. (2006). Methods for blind evaluation of noise variance in multichannel optical and radar images. *Telecommunications and Radioengineering*, Vol. 65 (6), pp. 509-537

Lukin V., Ponomarenko N., Abramov S., Vozel B., & Chehdi K. (2007). Improved noise parameter estimation and filtering of MM-band SLAR images. *Proceedings of MSMW 2007*, Vol. 1, pp. 439-441, Kharkov, Ukraine

Lukin V., Abramov S., Vozel B., Chehdi K., & Astola J. (2008a). Segmentation-based method for blind evaluation of noise variance in images. *SPIE Journal on Applied Remote Sensing*, Vol. 2, Aug. 2008, open access paper

Lukin V., Ponomarenko N., Abramov S., Vozel B., Chehdi K., & Astola J. (2008b). Filtering of radar images based on blind evaluation of noise characteristics. *Proceedings Image and Signal Processing for Remote Sensing XIV*, Cardiff, UK, September 2008, SPIE Vol. 7109

Lukin V., Abramov S., Ponomarenko N., Uss M., Vozel B., Chehdi K., & Astola J. (2009a). Processing of images based on blind evaluation of noise type and characteristics. *Proceedings of SPIE Symposium on Remote Sensing*, Vol. 7477, Berlin, Germany, September 2009

Lukin V., Krivenko S., Zriakhov M., Ponomarenko N., Abramov S., Kaarna A., & Egiazarian K. (2009b). Lossy compression of images corrupted by mixed Poisson and additive noise. *Proceedings of LNLA*, pp. 33-40, Helsinki, August 2009

Lukin V., Abramov S., Uss M., Vozel B., & Chehdi K. (2010a). Performance Analysis of Segmentation-based Method for Blind Evaluation of Additive Noise in Images. *Proceedings of MSMW*, Kharkov, Ukraine, June 2010

Lukin V., Fevralev D., Ponomarenko N., Abramov S., Pogrebnyak O., Egiazarian K., & Astola J. (2010b). Discrete cosine transform-based local adaptive filtering of images corrupted by nonstationary noise. *Electronic Imaging Journal*, 19(2), 1, April-June 2010

Lukin V., Abramov S., Ponomarenko N., Uss M., Zriakhov M., Vozel B., Chehdi K., & Astola J. (2011). Methods and Automatic Procedures for Processing Images Based on Blind Evaluation of Noise Type and Characteristics. *SPIE Journal on Advances in Remote Sensing*, DOI: 10.1117/1.3539768

Mallat S. (1998). *A Wavelet tour of signal processing,* Academic Press, San Diego

Murtagh F., Starck J.L., & Bijaoui A. (1995). Image restoration with noise suppression using a multiresolution support, *Astron. Astrophys. Suppl. Ser*, 112, pp. 179-189

Öktem R. & Egiazarian K. (1999). Transform Domain Algorithm for Reducing the Effect of Film-Grain Noise in Image Compression. *Electronic Letters*, Vol. 35, No. 21, pp. 1830-1831

Oliver C. & Quegan S. (2004). *Understanding Synthetic Aperture Radar Images*, SciTech Publishing

Petrovic N., Zlokolica V., Goossens B., Pizurica A., & Philips W. (2007). Characterization of correlated noise in video sequences and its applications to noise removal. *Proceedings of the 3rd International Workshop on Video Processing and Quality Metrics for Consumer Eletronics*, Scottsdale, Arizona, USA, January 2007

Plataniotis K.N. & Venetsanopoulos A.N. (2000). *Color Image Processing and Applications*, Springer-Verlag, NY

Ponomarenko N., Lukin V., Egiazarian K., & Lepisto L. (2011). Color image lossy compression based on blind evaluation and prediction of noise characteristics. Accepted to *SPIE Conference Image Processing: Algorithms and Systems VII*, Vol. 7870, 12 p., San Francisco, USA

Pratt W.K. (2007). *Digital Image Processing* (Fourth Edition). Wiley-Interscience, NY, USA

Ramponi G. & D'Alvise R. (1999). Automatic Estimation of the Noise Variance in SAR Images for Use in Speckle Filtering, *Proceedings of EEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, Vol. 2, pp. 835-838, Antalya, Turkey

Sendur L. & Selesnick I.W. (2002). Bivariate shrinkage with local variance estimation. *IEEE Signal Processing Letters*, Vol. 9, No. 12, pp. 438-441

Solbo S. & Eltoft T. (2004). Homomorphic Wavelet-based Statistical Despeckling of SAR Images. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. GRS-42, No. 4, pp. 711-721

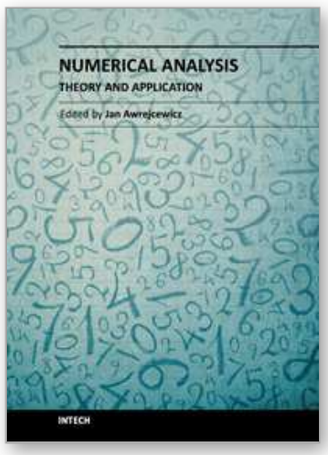
Touzi R. (2002). A Review of Speckle Filtering in the Context of Estimation Theory. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, No. 11, pp. 2392-2404

Uss M., Vozel B., Lukin V., & Chehdi K. (2011). Local Signal-Dependent Noise Variance Estimation from Hyperspectral Textural Images. *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 2, DOI: 10.1109/JSTSP.2010.2104312

Vozel B., Chehdi K., & Klaine L. (2006). Noise identification and estimation of its statistical parameters by using unsupervised variational classification. *Proceedings of ICASSP*, Vol. II, pp. 841-844

Vozel B., Abramov S., Chehdi K., Lukin V., Ponomarenko N., Uss M., & Astola J. (2009). Blind methods for noise evaluation in multi-component images, In: *Multivariate Image Processing,* pp. 263-295, France

Yu-jin Zhang (January 2006). *Advances in Image and Video Segmentation*, IRM Press, ISBN 1591407559



Numerical Analysis - Theory and Application

Edited by Prof. Jan Awrejcewicz

ISBN 978-953-307-389-7
Hard cover, 626 pages
Publisher InTech
Published online 09, September, 2011
Published in print edition September, 2011

Numerical Analysis â€“ Theory and Application is an edited book divided into two parts: Part I devoted to Theory, and Part II dealing with Application. The presented book is focused on introducing theoretical approaches of numerical analysis as well as applications of various numerical methods to either study or solving numerous theoretical and engineering problems. Since a large number of pure theoretical research is proposed as well as a large amount of applications oriented numerical simulation results are given, the book can be useful for both theoretical and applied research aimed on numerical simulations. In addition, in many cases the presented approaches can be applied directly either by theoreticians or engineers.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Sergey Abramov, Victoria Zabrodina, Vladimir Lukin, Benoit Vozel, Kacem Chehdi and Jaakko Astola (2011). Methods for Blind Estimation of the Variance of Mixed Noise and Their Performance Analysis, Numerical Analysis - Theory and Application, Prof. Jan Awrejcewicz (Ed.), ISBN: 978-953-307-389-7, InTech, Available from: http://www.intechopen.com/books/numerical-analysis-theory-and-application/methods-for-blind-estimation-of-the-variance-of-mixed-noise-and-their-performance-analysis

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen