Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers

Bruno Aiazzi ([1])**, Luciano Alparone** ([2])**, Alessandro Barducci** ([1])**, Stefano Baronti** ([1])**, Paolo Marcoionni** ([1])**,
Ivan Pippi** ([1]) **and Massimo Selva** ([1])

([1]) *Istituto di Fisica Applicata «Nello Carrara» (IFAC), CNR, Firenze, Italy*
([2]) *Dipartimento di Elettronica e Telecomunicazioni, Università degli Studi di Firenze, Italy*

Abstract
The definition of noise models suitable for hyperspectral data is slightly different depending on whether whisk-broom or push-broom are dealt with. Focussing on the latter type (*e.g.*, VIRS-200) the noise is intrinsically non-stationary in the raw digital counts. After calibration, *i.e.* removing the variability effects due to different gains and offsets of detectors, the noise will exhibit stationary statistics, at least spatially. Hence, separable 3D processes correlated across track (x), along track (y) and in the wavelength (λ), modelled as auto-regressive with GG statistics have been found to be adequate. Estimation of model parameters from the true data is accomplished through robust techniques relying on linear regressions calculated on scatter-plots of local statistics. An original procedure was devised to detect areas within the scatter-plot corresponding to statistically homogeneous pixels. Results on VIRS-200 data show that the noise is heavy-tailed (tails longer than those of a Gaussian PDF) and somewhat correlated along and across track by slightly different extents. Spectral correlation has been investigated as well and found to depend both on the sparseness (spectral sampling) and on the wavelength values of the bands that have been selected.

1. Introduction

Airborne imaging spectrometers exhibit characteristics somewhat different from those of conventional multi-spectral scanners (Landsat TM, SPOT, etc.). In particular, noise processes may be non-Gaussian, due to non-linearity of sensors and to pre-processing of raw data, and correlated, both spatially and spectrally, because of the mechanism of opto-electronic scanning. Hence, most analysis methods, conceived to effectively operate in the presence of Additive White Gaussian Noise (AWGN), may become inadequate to the present case. By exploiting a Probability Density Function (PDF) capable of representing also non-Gaussian distributions, auto-correlated non-Gaussian noise models may be described as a generalisation of auto-regressive models. The noise parameters, specifically variance, Correlation Coefficients (CCs), and shape factor of the Generalised Gaussian (GG) PDF are straightforwardly estimated from the observed data by means of robust statistical procedures based on multivariate regressions.

Concerning exploitation of noise analysis methods in application fields, the knowledge of the parametric noise model may be used to develop data-dependent spectral decompositions,

either orthogonal (Principal Components Analysis-PCA), or non-orthogonal, like Projection Pursuit (PP) (Jimenez and Landgrebe, 1999) matched to the noise. Given the analytical, as well as statistical, intractability of highly dimensional data (Jimenez and Landgrebe, 1998), a reduction in dimension of the data space may be achieved by projecting hyperspectral pixel vectors onto a sub-space orthogonal to that of the noise, like in the Noise-Adjusted PCA (NAPCA) (Lee *et al.*, 1990).

Spectral decompositions often represent a necessary pre-processing step in object detection applications. In fact, they may be used to remove the spatially variant signal due to the scene background (background removal). Such a step simplifies the derivation of the detection algorithm in that it can be designed to operate over the residual clutter that can be considered (almost) spatially stationary. Noise reduces the effectiveness of these techniques in a manner that depends both on the statistical distribution of its intensity and on its correlation properties. In this context, noise models whose parameters are adaptively tuned to the image sequences represent a valuable tool to analyze the effects of noise on the performance expected from background removal algorithms.

The parametric noise model is also useful to estimate the theoretical entropy bounds of reversible hyperspectral data compression methods, as well as to measure the spectral information content (Aiazzi *et al.*, 2001), *i.e.* the information the sensor would yield in the case of an ideally noise-free process of sensing and digitisation of spectral radiance.

Experimental results of noise modelling and estimation, carried out both on simulated noisy images and on data acquired by the Visible InfraRed Scanner (VIRS-200), are reported and discussed to corroborate the assumptions underlying the proposed model and to validate the calibration procedure (Barducci and Pippi, 2001) undergone by the raw data.

2. Noise modelling

This section will focus on estimation of noise variance from true images. Unlike coherent or systematic disturbances, which may occur in

some kind of sensed data, the noise is assumed to be due to a fully stochastic process. Solutions will be devised for estimating the noise typically introduced by multispectral scanners and especially by hyperspectral imaging spectrometers. Let us assume for the noise an additive signal-independent model

$$g(i, j, k) = f(i, j, k) + n(i, j, k) \qquad (2.1)$$

in which $g(i, j, k)$ is the recorded intensity at pixel position (i, j) of the kth spectral band and $f(i, j, k)$ the kth component of the spectral reflectance at (i, j). Both $g(i, j, k)$ and $f(i, j, k)$ are regarded as non-stationary non-Gaussian (auto)correlated stochastic processes. The term $n(i, j, k)$ is a zero-mean Gaussian process independent of f, stationary along (i, j) but not along k and correlated, both spatially and spectrally. The variance is $\sigma_n^2(k)$, and the CC's are ρ_x and ρ_y, across and along the track, respectively, and ρ_λ in the spectral direction. The CCs are assumed to be constant throughout.

The variance of (2.1) can be easily calculated as

$$\sigma_g^2(i, j, k) = \sigma_f^2(i, j, k) + \sigma_n^2(k) \qquad (2.2)$$

thanks to the independence between radiance and noise components and to the spatial stationarity of the latter. From (2.2) it appears that $\sigma_n^2(k)$ can be estimated by averaging $\sigma_g^2(i, j, k)$ in homogeneous areas of the kth band, where $\sigma_f^2(i, j, k) \equiv 0$, by definition. The homogeneity requirement is crucial because the radiance components are generally correlated with one another, besides being spatially autocorrelated. Since an area may be homogeneous in one band, but not in another, in the following the analysis will be carried out on separate bands, notwithstanding they may exhibit noise components that are dependent on one another. For sake of clarity the index k will be omitted to indicate a generic band.

3. Estimation of noise variance and correlation

The standard deviation of the noisy observed band $g(i, j)$ may be stated in homogeneous areas as

$$\sigma_g(i,j) = \sigma_n. \qquad (3.1)$$

Therefore, (3.1) yields an estimate of σ_n, namely $\hat{\sigma}_n$, as the y-intercept of the horizontal regression line drawn on the scatter-plot of $\hat{\sigma}_g$ *versus* $\hat{\mu}_g$, in which the symbol ˆ denotes estimated values, and calculated only on pixels belonging to homogeneous areas. Although methods based on scatter-plots were devised more than one decade ago for speckle noise assessment (Lee and Hoppel, 1989), the crucial point is how to reliably identify homogeneous areas. To overcome the drawback of a usersupervised method, an automatic procedure was developed (Aiazzi *et al.*, 1999a), based on the fact that each homogeneous area originates a cluster of scatter-points. All these clusters are aligned along a horizontal straight line having y-intercept equal to σ_n. Instead, the presence of signal edges and textures originates scatter-points spread throughout the plot.

The scatter-plot relative to the whole band may be regarded as the joint Probability Density Function (PDF) of estimated local standard deviation to estimated local mean. In the absence of any signal textures, *e.g.*, the image is made up of uniform noisy patches, by assuming that the noise is stationary, the PDF will be given by the superposition of as many unimodal distributions as there are patches. Since the noise is independent of the signal, the measured variance does not depend on the underlying mean. Thus, all the above expectations will be aligned along a horizontal line.

The presence of textured areas will modify the «flaps» of the PDF, which will still exhibit aligned modes, or possibly a watershed. The idea is thresholding the PDF to identify a number of points belonging to the most homogeneous image areas, large enough to yield a statistically consistent estimate, and small enough to avoid comprising signal textures that, even if weak, might introduce a bias by excess.

Now let us calculate the space-varying (auto) covariance of unity lag along either of the coordinate directions, say i,

$$C_g(i,j;\,1,\,0) \triangleq E\{[g(i,j)-E(g(i,j))]\cdot[g(i+1,j)+ \\ -E(g(i+1,j))]\} = C_f(i,j;\,1,\,0)+\rho_x\cdot\sigma_n^2. \qquad (3.2)$$

The term $C_f(i,j;\,1,\,0)$ on right-hand of (3.2), is

identically zero in homogeneous areas, in which $\sigma_g(i,j)$ becomes equal to σ_n. Thus, (3.2) becomes

$$C_g(i,j;\,1,0)=\rho_x\cdot\sigma_n^2=\rho_x\cdot\sigma_g(i,j)\cdot\sigma_g(i+1,j). \quad (3.3)$$

Hence, ρ_x, and analogously ρ_y, may be estimated from those points, lying on the covariance-to-variance scatter-plots, corresponding to homogeneous areas.

To avoid calculating the PDF, the following procedure was devised:

1. Within a $(2m+1)\times(2m+1)$ window sliding over the image calculate the local statistics of the noisy image to estimate its space-varying ensemble statistics:

– Average $\bar{g}(i,j) \equiv \hat{\mu}_g(i,j)$

$$\bar{g}(i,j) = \frac{1}{(2m+1)^2} \sum_{k=-m}^{m} \sum_{l=-m}^{m} g(i+k,j+l). \,(3.4)$$

– RMS deviation from the average, $\hat{\sigma}_g(i,j)$, with

$$\hat{g}_g^2(i,j) = \frac{1}{(2m+1)^2-1} \cdot \\ \cdot \sum_{k=-m}^{m} \sum_{l=-m}^{m} [g(i+k,j+l) - \bar{g}(i,j)]^2. \qquad (3.5)$$

– Cross-deviation from the average $\hat{C}_g(i,j;1,0)$, given by

$$\hat{C}_g(i,j;1,0) = \frac{1}{(2m+1)^2-1} \cdot \\ \cdot \sum_{k=-m}^{m} \sum_{l=-m}^{m} [g(i+k,j+l) - \bar{g}(i,j)] \cdot \\ \cdot [g(i+k+1,j+l) - \bar{g}(i+1,j)]. \qquad (3.6)$$

2. Draw the scatter-plots either of $\hat{\sigma}_g(i,j)$ *versus* $\hat{\mu}_g(i,j)$, to estimate σ_n, or of $\hat{C}_g(i,j;\,1,\,0)$ *versus* $\hat{\sigma}_g^2(i,j)\cdot\hat{\sigma}_g^2(i+1,j)$, to estimate ρ_x.

3. Partition the scatter-plot planes into an $L\times L$ array of rectangular blocks.

4. Sort and label such blocks by decreasing population, *i.e.* number of scatter-points: if $\mathcal{C}(\cdot)$ denotes the cardinality of a set, then $\mathcal{C}(B^{(k)}) \geq \\ \geq \mathcal{C}(B^{(k+1)})$, $k=1,...,L^2$.

5. For each scatter-plot, calculate a succession of horizontal regression lines, that is $\{\hat{\sigma}_n(k)\}$, from the set of scatter points $\left\{ \boldsymbol{p} \,|\, \boldsymbol{p} \in \bigcup_{l=1}^{k} B^{(l)} \right\}$.

The succession attains a steady value of the parameter after a number of terms that depends on the actual percentage of homogeneous points. The size of the partition and the stop criterion are non-crucial: a 100×100 array of blocks and stop after processing $2 \div 10\%$ of points, depending on the degree of inhomogeneity of the scene. The size of the processing window, *i.e.* $(2m+1) \times \times (2m+1)$, is non-crucial if the noise is white. Otherwise a size 7×7 to 11×11 is recommended, since too small a size will underestimate the covariance (Aiazzi *et al.*, 1999b).

It is noteworthy that, unlike most methods which rely on the assumption of white noise, the scatter-plot methods, besides being easily adjustable to deal with signal-dependent noise, can accurately measure also correlated noise, and thus have been preferred in this context.

4. Estimation of noise distribution

4.1. *White noise*

Once the most homogeneous image pixels have been made available by the scatter-plot procedure, the PDF of the noise may be estimated (Aiazzi *et al.*, 2002) by calculating the histogram of the following amount:

$$\hat{n}(i,j) = \sqrt{\frac{(2m+1)^2}{(2m+1)^2 - 1}} \cdot [g(i,j) - \bar{g}(i,j)]$$

(4.1)

only on those pixels recognized to be homogeneous. The correcting factor under square root in (4.1) is easily calculated by imposing that the PDF has variance σ_n^2, since $g - \bar{g}$ yields an estimate biased by defect of the noise variance, due to the statistical dependence of the noise components affecting $g(i, j)$ and $\bar{g}(i, j)$.

4.2. *Correlated noise*

When the noise is spatially correlated (4.1) no longer holds, because of the statistical dependence of $n(i)$ from the previous noise samples. Let us assume for the stationary zero-mean Gaussian noise a separable first-order

Markov model, uniquely defined, in one dimension, by the ρ and the σ_n^2

$$n(i) = \rho \cdot n(i-1) + \varepsilon_n(i) \qquad (4.2)$$

in which $\varepsilon_n(i)$ is a white Gaussian random process having variance $\sigma_{\varepsilon n}^2 = \sigma_n^2 \cdot (1 - \rho^2)$. From (4.2) it stems that the noise covariance of lag m is

$$E[n(i) \cdot n(i+m)] = \rho^{|m|} \cdot \sigma_n^2. \qquad (4.3)$$

If (4.2) and (4.3) are utilized to calculate the correlation of the noise affecting g and $\bar{g}$ on a homogeneous window, the estimated noise at pixel (i, j) can be written as

$$\hat{n}(i,j) =$$

$$= \sqrt{\frac{(2m+1)^2}{(2m+1)^2 - \left(1 + 2\rho_x \dfrac{1-\rho_x^m}{1-\rho_x}\right)\left(1 + 2\rho_y \dfrac{1-\rho_y^m}{1-\rho_y}\right)}} \cdot$$

$$\cdot [g(i,j) - \bar{g}(i,j)] \qquad (4.4)$$

in which $2m+1$ is the length of the sliding window. Notice that (4.4) reduces to (4.1) when the CCs are both zero.

4.3. *Generalised Gaussian PDF*

A model suitable for describing non-Gaussian noise may be achieved by varying the parameters ν (shape factor) and σ (standard deviation) of the Generalised Gaussian Density (GGD), which is defined as

$$p(x) = \left[\frac{\nu \cdot \eta(\nu,\sigma)}{2 \cdot \Gamma(1/\nu)}\right] \exp\left\{-[\eta(\nu,\sigma) \cdot |x|]^\nu\right\}$$

(4.5)

in which $\eta(\nu,\sigma) = \dfrac{1}{\sigma}[\Gamma(3/\nu)/\Gamma(1/\nu)]^{1/2}$ (4.6)

and $\Gamma(\cdot)$ is the Gamma function, *i.e.* $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$, $z > 0$. Since $\Gamma(n) = (n-1)!$, when $\nu = 1$ a Laplacian Law is obtained, while $\nu = 2$ yields a Gaussian distribution. As limit cases, for $\nu \to 0$, $p(x)$ becomes an impulse function, yet having broad and flat tails and thus nonzero σ^2 variance, whereas for $\nu \to \infty$, $p(x)$ approach-
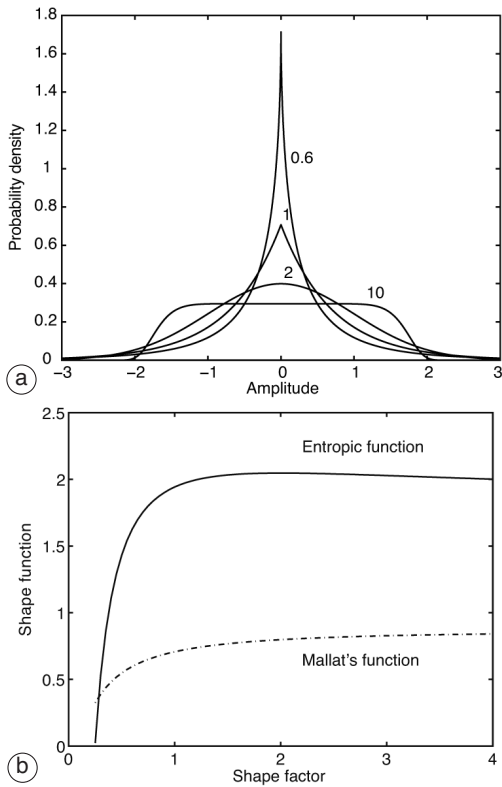
Fig. 1a,b. a) unity-variance GG function plotted for several ν's; b) entropy function and Mallat's function of a unity-variance GG PDF as a function of the shape factor ν.

es a uniform distribution having variance σ^2 as well. The shape parameter ν rules the exponential rate of decay: the larger the ν the flatter the PDF; the smaller the ν, the more peaked the PDF. Figure 1a shows the trend of the GG function for different values of ν.

The matching between a GGD and the empirical data can be obtained either through an original method, recently developed by the authors (Aiazzi *et al.*, 1999c), based on fitting the entropy of the modelled source to that of the empirical data, or by the simple method introduced by Mallat in 1989 and reinvented six years later (Sharifi and Leon-Garcia, 1995). This method substantially falls into the category of moment methods, being based on matching the moments of the data set

with those of the assumed distribution, in order to find out its variance and shape factor. For a GGD, the ratio of mean absolute value to standard deviation is a steadily increasing function of the ν

$$\mathcal{F}_M(\nu) = \frac{\Gamma(2/\nu)}{\sqrt{\Gamma(1/\nu) \cdot \Gamma(3/\nu)}}. \qquad (4.7)$$

Given a collection of N *i.i.d.* zero-mean random variables, $\{x_1, x_2, ..., x_N\}$, following a GGD $p(x)$, let $m_1 = 1/N \sum_{i=1}^{N} |x_i|$ be the estimate of the mean absolute value and $\hat{\sigma}^2 = 1/N \sum_{i=1}^{N} x_i^2$ the sample variance. The parameter ν is estimated by inverting (4.7), that is by solving

$$\hat{\nu} = \mathcal{F}_M^{-1}\left(\frac{m_1}{\hat{\sigma}}\right). \qquad (4.8)$$

The entropy method (Aiazzi *et al.*, 1999c) cannot be utilized around $\nu = 2$ because the entropy function takes its maximum in the Gaussian case. Figure 1b shows the entropy function and Mallat's function plotted *versus* the shape factor ν. As it appears, inversion of the former is not possible around $\nu = 2$; instead its slope is greater than that of Mallat's function for $\nu < 1$. Hence, the entropy method is more accurate than Mallat's method for $\nu < 1$.
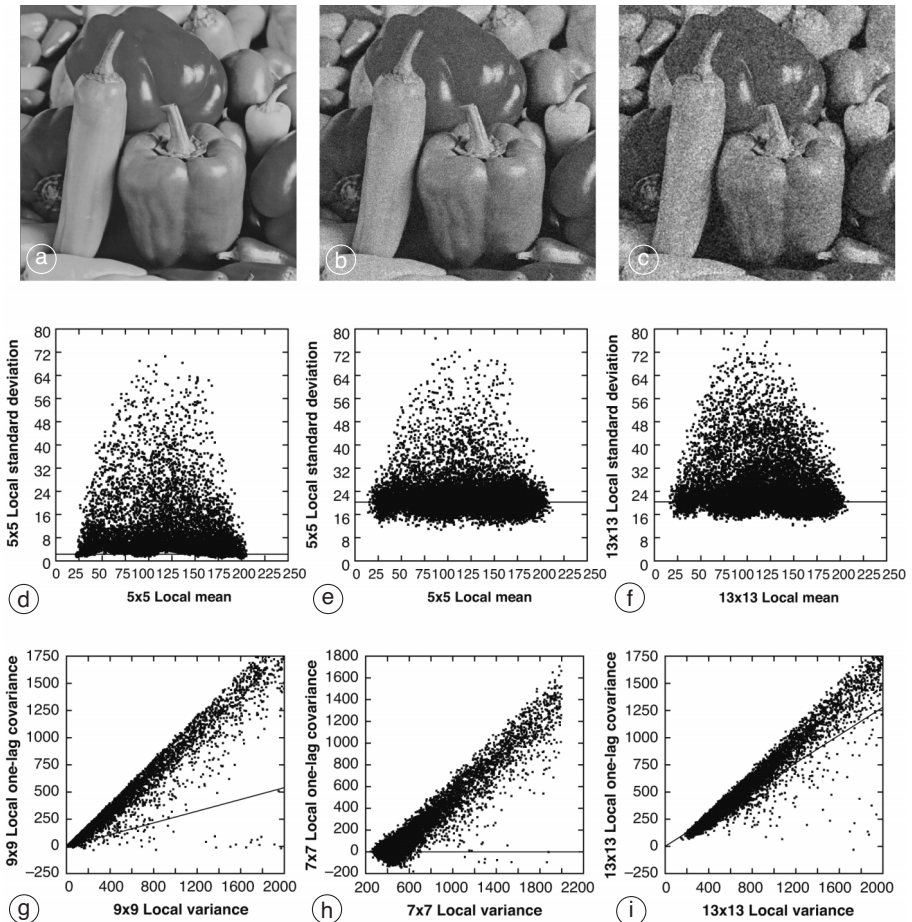
5. Experimental results

5.1. *Simulated data*

The procedures for information assessment described in Aiazzi *et al.* (2001) will be firstly validated on true image data corrupted with simulated additive zero-mean Gaussian noise, both white and spatially correlated. Zero-mean Gaussian noise with $\sigma_n^2 = 400$, either with $\rho_x = \rho_y = 0$, or with $\rho_x = \rho_y = 0.65$, was superimposed to a grey-scale 8-bit image, referred to as Peppers. The test image and its noisy versions are portrayed in fig. 2a-i. Table I summarizes the results obtained.

5.2. *VIRS-200 data*

The data set comprises a hyperspectral image collected by the Visible InfraRed Scanner

Bruno Aiazzi, Luciano Alparone, Alessandro Barducci, Stefano Baronti, Paolo Marcoionni, Ivan Pippi and Massimo Selva

Fig. 2a-i. a) Test image Peppers and its noisy versions; b) additive white zero-mean Gaussian noise with variance $\sigma_n^2 = 400$; c) additive correlated Gaussian noise with $\sigma_n^2 = 400$ and $\rho_x = \rho_y = 0.65$. Scatter-plots of $\hat{\sigma}_g$ *versus* $\hat{\mu}_g$ (with regression lines superimposed) for d) original Peppers, e) Peppers with white zero-mean Gaussian noise ($\sigma_n^2 = 400$), and f) Peppers with correlated zero-mean Gaussian noise ($\sigma_n^2 = 400$ and $\rho_x = \rho_y = 0.65$). Scatter-plots of $\hat{C}_g$ *versus* $\hat{\sigma}_g^2$ (with regression lines superimposed) for g) original Peppers, h) additive white zero-mean Gaussian noise with $\sigma_n^2 = 400$; i) additive correlated zero-mean Gaussian noise with $\sigma_n^2 = 400$ and $\rho_x = \rho_y = 0.65$. All regression lines are calculated on subsets of scatter-points recognized as belonging to homogeneous areas.

Table I. True and estimated noise parameters for original and simulated noisy versions, white and correlated, of the test image Peppers.

	σ_u	$\hat{\sigma}_u$	ρ_x	$\hat{\rho}_x$	ρ_y	$\hat{\rho}_y$	ν	$\hat{\nu}$
Original	N/A	2.5	N/A	0.27	N/A	0.11	N/A	1.81
White noise	20.1	20.0	0	0.02	0	0.02	2	2.07
Correlated noise	20.1	20.3	0.65	0.62	0.65	0.63	2	2.12

(VIRS-200), along the coast of Tuscany, in Central Italy. The image is constituted by 20 narrow bands (2.5 nm) out of 240 recorded at different wavelengths, programmable in the range 400÷1000 nm. The size of each image is 512 pixels across track. The raw data were originally 12 bit. All the data have been radiometrically calibrated, and thus expressed as radiance values, and packed in a 16-bit word. De-striping is particularly crucial (Barducci and Pippi, 2001), since the sensors in the array moved along track exhibit different offsets and gains at each wavelength. A couple of bands, one in the visible (VIS), another in the near-infrared (NIR) wavelengths are shown in fig. 3a,b.

The CCs of the noise were estimated by means of scatter-plots of local statistics calculated on 11×11 windows. This relatively large window size is dictated by the requirement of ergodicity, which allows unbiased statistics to be estimated from a single realisation of the random process. CCs are underestimated if the window is too small, because a large correlation may not be captured by a small window. On the
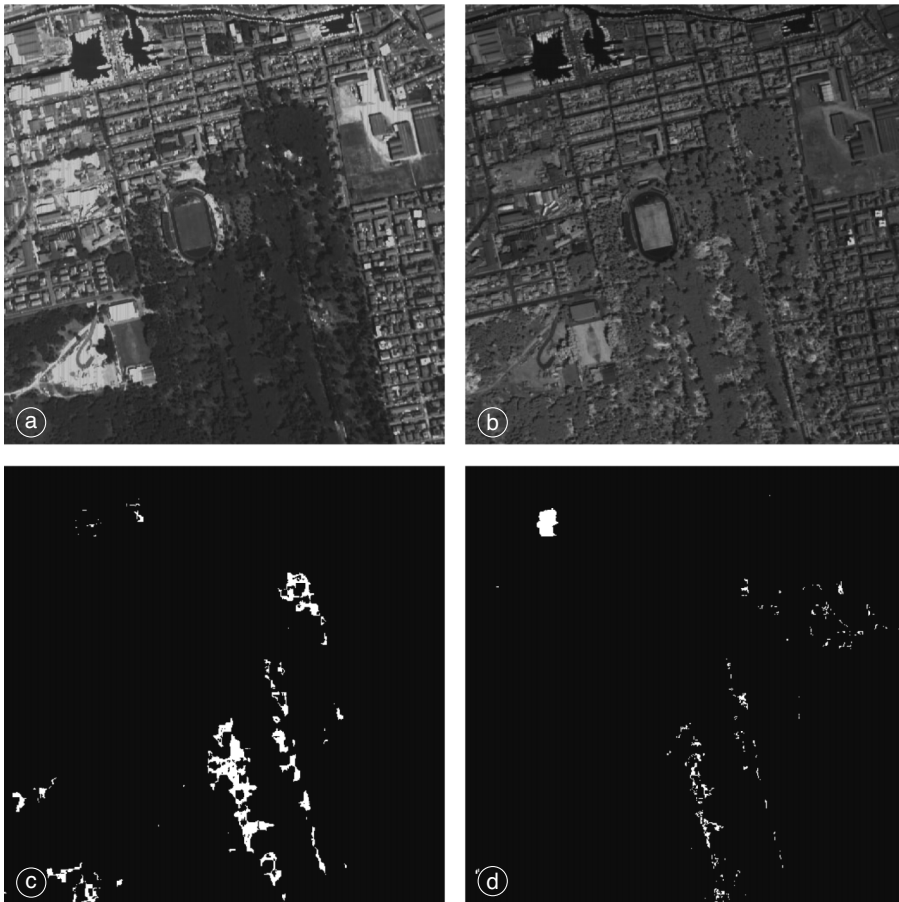

Fig. 3a-d. 512×512 details taken from two bands of the VIRS-200 image of the city of Viareggio: a) VIS wavelengths (554 nm); b) NIR wavelengths (889 nm); c) homogeneous pixels of band 8 (2% of total); d) homogeneous pixels of band 18 (0.6% of total).

other side, the window size should not be too large, to prevent textured pixels from being considered in the analysis.

A 2% of points was utilized in the VIS bands; 0.6% for the NIR bands, since a larger amount of pixels would inject textures (inhomogeneities) into the calculation of the PDFs. The binary mask of selected points is shown in fig. 3c,d, for the VIS and NIR band, respectively. The CCs measured are $\hat{\rho}_x=0.50$ and $\hat{\rho}_y=0.70$ for the VIS band; $\hat{\rho}_x=0.55$ and $\hat{\rho}_y=0.65$ for the NIR band. Similar values were found for the remaining 18 bands. Spectral correlation has been investigated as well and found to depend both on the sparseness (spectral sampling) and on the wavelength values of the bands that have been selected. Results are not reported here both for sake of conciseness, and as irrelevant for intra-band noise analysis.

The noise PDFs were calculated for the two sample bands and plotted in fig. 4a,b together with generalised Gaussian PDFs having same variance and shape factor as the distribution of the noise samples estimated by means of (4.4). As it appears, the degree of matching of the theoretical and empirical distributions is impressive. The VIS band exhibits a hyper-Gaussian PDF ($v=1.57$), while the NIR band is approximately Laplacian ($v=1.18$). In both cases tails are heavier than those of a Gaussian PDF ($v=2$).
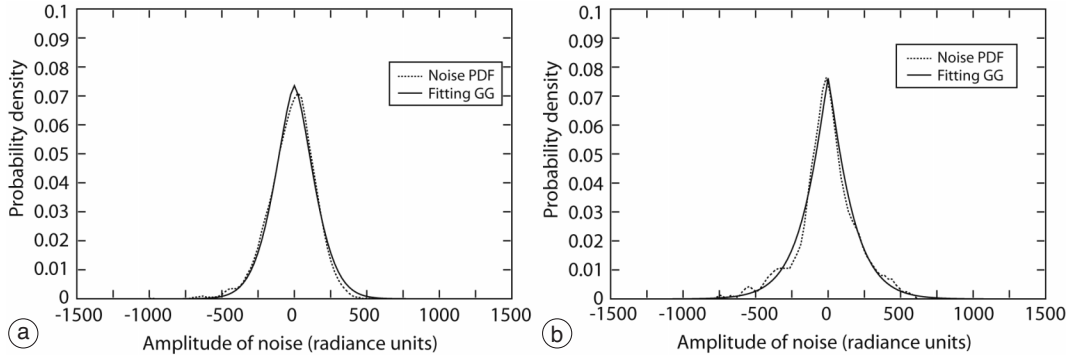
Fig. 4a,b. Normalised noise histograms and fitting GG PDFs: a) 554 nm, $\sigma=158$, $v=1.57$; b) 889 nm, $\sigma=205$, $v=1.18$.
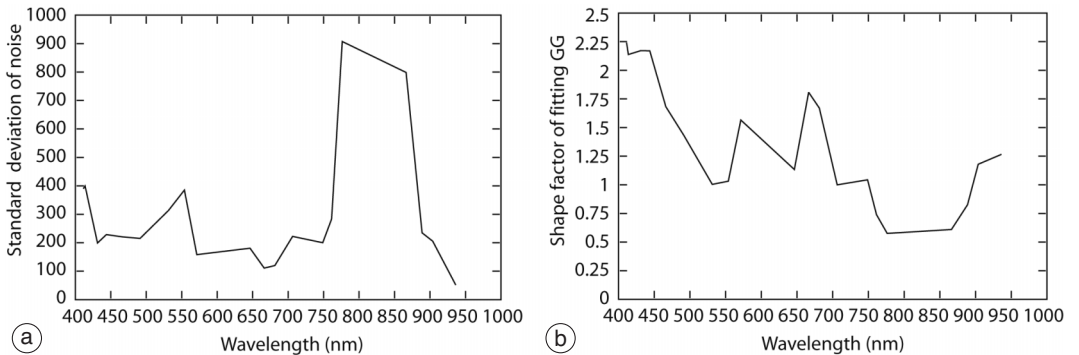
Fig. 5a,b. Estimated parameters of noise PDF of the VIRS-200 image *versus* wavelength: a) standard deviation $\hat{\sigma}_n$; b) shape factor $\hat{v}$.

The noise variance is slightly higher in the NIR wavelengths than in the VIS wave-lengths, as expected from theory.

After calibration, also aimed at removing the variability effects due to different gains and offsets of detectors, the noise exhibits statistics stationary in space, but not spectrally. Therefore, an analysis varying with wavelength of noise variance and shape factor was carried out. Figure 5a,b highlights how the two parameters, standard deviation and shape factor, change with the wavelength. The fact that noise tails becomes heavier as the wavelength increases is peculiar of push-broom scanners and mainly depends on the de-striping procedure. Ideally, Gaussian PDFs should be obtained, same as it happens for single-detector whisk-broom spectrometers, like AVIRIS (Aiazzi *et al.*, 2001). The non-ideal equalisation of detectors' gains makes the noise to be slightly non-stationary. Hence, the superposition of many Gaussian PDFs having variances different from one another, originates a global PDF having tails longer than those of a Gaussian density.

6. Conclusions and developments

A robust procedure for noise estimation in a sequence of hyperspectral bands has been described and assessed. Noise PDFs and CCs in each spectral band are automatically calculated from the data. Knowledge of the noise PDF, especially if not heavy-tailed, may help exactly circumscribe the range of noise values, with the benefit that the loss of performance expected on application tasks and due to the noise can be accurately modelled and hence predicted to a large extent.

The procedure described in this paper could be profitably utilized to assess the calibration accuracy of the data, especially for de-striping, whose performance can be measured as Gaussianity of the noise in each spectral band. Eventually, spectral correlation of the noise, which is low for sparse bands, but can significantly grow when hundredths of bands are delivered by the on-board instrument, can be measured by devising a further bivariate regression between

homogeneous pixels in a couple of consecutive bands.

REFERENCES

AIAZZI, B., L. ALPARONE and S. BARONTI (1999a): Reliably estimating the speckle noise from SAR data, in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 1546-1548.

AIAZZI, B., L. ALPARONE and S. BARONTI (1999b): Unsupervised assessment and pyramidal filtering of colored speckle, in *SAR Image Analysis, Modelling, and Techniques*, IV, edited by F. POSA, *SPIE Proc.*, *EUROPTO Ser.*, **3869**, 9-20.

AIAZZI, B., L. ALPARONE and S. BARONTI (1999c): Estimation based on entropy matching for generalized Gaussian PDF modeling, *IEEE Signal Processing Lett.*, **6**, 138-140.

AIAZZI, B., L. ALPARONE, A. BARDUCCI, S. BARONTI and I. PIPPI (2001): Information-theoretic assessment of sampled hyperspectral imagers, *IEEE Trans. Geosci. Remote Sensing*, **39**, 1447-1458.

AIAZZI, B. L. ALPARONE, A. BARDUCCI, S. BARONTI and I. PIPPI (2002): Estimating noise and information of multispectral imagery, *J. Optical Engin.*, **41**, 656-668.

BARDUCCI, A. and I. PIPPI (2001): Analysis and rejection of systematic disturbances in hyperspectral remotely sensed images of the Earth, *Appl. Optics*, **40**, 1464-1477.

JIMENEZ, L.O. and D.A. LANDGREBE (1998): Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data, *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, **28**, 39-54.

JIMENEZ, L.O. and D.A. LANDGREBE (1999): Hyperspectral data analysis and supervised feature reduction with projection pursuit, *IEEE Trans. Geosci. Remote Sensing*, **37**, 2653-2667.

LEE, J.-S. and K. HOPPEL (1989): Noise modeling and estimation of remotely sensed images, in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 1005-1008.

LEE, J.B., A.S. WOODYATT and M. BERMAN (1990): Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform, *IEEE Trans. Geosci. Remote Sensing*, **28**, 295-304.

SHARIFI, K. and A. LEON-GARCIA (1995): Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video, *IEEE Trans. Circuits Syst. Video Technol.*, **5** (1), 52-56.