Lossy compression of hyperspectral images based on noise parameters estimation and variance stabilizing transform

Alexander N. Zemliachenko
Ruslan A. Kozhemiakin
Mikhail L. Uss
Sergey K. Abramov
Nikolay N. Ponomarenko
Vladimir V. Lukin
Benoît Vozel
Kacem Chehdi

SPIE.

Lossy compression of hyperspectral images based on noise parameters estimation and variance stabilizing transform

Alexander N. Zemliachenko,[a] Ruslan A. Kozhemiakin,[a] Mikhail L. Uss,[a] Sergey K. Abramov,[a] Nikolay N. Ponomarenko,[a] Vladimir V. Lukin,[a] Benoît Vozel,[b,*] and Kacem Chehdi[b]

[a]National Aerospace University, Kharkov 61070, Ukraine
[b]University of Rennes 1, IETR UMR CNRS 6164, Lannion cedex 22305, France

Abstract. A problem of lossy compression of hyperspectral images is considered. A specific aspect is that we assume a signal-dependent model of noise for data acquired by new generation sensors. Moreover, a signal-dependent component of the noise is assumed dominant compared to a signal-independent noise component. Sub-band (component-wise) lossy compression is studied first, and it is demonstrated that optimal operation point (OOP) can exist. For such OOP, the mean square error between compressed and noise-free images attains global or, at least, local minimum, i.e., a good effect of noise removal (filtering) is reached. In practice, we show how compression in the neighborhood of OOP can be carried out, when a noise-free image is not available. Two approaches for reaching this goal are studied. First, lossy compression directly applied to the original data is considered. According to another approach, lossy compression is applied to images after direct variance stabilizing transform (VST) with properly adjusted parameters. Inverse VST has to be performed only after data decompression. It is shown that the second approach has certain advantages. One of them is that the quantization step for a coder can be set the same for all sub-band images. This offers favorable prerequisites for applying three-dimensional (3-D) methods of lossy compression for sub-band images combined into groups after VST. Two approaches to 3-D compression, based on the discrete cosine transform, are proposed and studied. A first approach presumes obtaining the reference and "difference" images for each group. A second performs compression directly for sub-images in a group. We show that it is a good choice to have 16 sub-images in each group. The abovementioned approaches are tested for Hyperion hyperspectral data. It is demonstrated that the compression ratio of about 15–20 can be provided for hyperspectral image compression in the neighborhood of OOP for 3-D coders, which is sufficiently larger than for component-wise compression and lossless coding.

1 Introduction

Airborne and space-borne hyperspectral imaging is now a useful and conventional tool in remote sensing (RS). It is able to provide valuable information for different applications, such as Earth surface monitoring, oceanology and hydrology, pollution control, and so on.[1–3] However, taking into account modern tendencies of increasing spectral and spatial resolution of sensors as well as areas of sensed terrains, the amount of acquired data greatly increases. Then, there is often an

*Address all correspondence to: Benoit Vozel, E-mail: benoit.vozel@univ-rennes1.fr

urgent need to compress the hyperspectral data intended for transferring, archiving, and dissemination.[3–6]

As it is known, there exist lossless and lossy data compression techniques.[4,7] Some customers and researchers consider lossless compression to be preferable.[6] However, lossless compression techniques applicable to hyperspectral data have a serious drawback. Even the best among them provide a compression ratio (CR) of about 4–5, [5,6,8] and there are no recent papers demonstrating considerable improvement of CR for lossless compression techniques. Meanwhile, it is often desirable or required to have a larger CR. This means that one has to apply the lossy compression.

Two aspects are worth stating here. First, lossy compression can be the only way out in some practical situations. Second, under certain conditions, useful information can be practically not lost in images if one applies lossy compression. This happens if distortions introduced by lossy compression are smaller than the level of the noise present in an image, and these distortions basically relate to noise removal.[9,10] Thus, it is possible to employ near-lossy compression approaches[10] and to sufficiently provide larger attained CR values due to this.

It is also worth mentioning that a partial noise filtering effect that takes place in lossy compression of noisy images can be considered as a positive factor. It can lead to improving the classification accuracy of compressed hyperspectral images compared to uncompressed (original, compressed in a lossless manner) images.[11] It can even lead to a better visual quality of compressed images compared to original (noisy) ones if the parameters of a coder are properly adjusted and noise level in the original images is quite high.[12] A question, then, is what is "properly adjusted" with respect to properties of hyperspectral RS data and noise present in the corresponding images.

For many years, noise in acquired hyperspectral data has been assumed additive,[10,13] and this is almost true for images provided by such hyperspectral sensors as airborne visible/infrared imaging spectrometer (AVIRIS). However, studies performed recently have clearly shown that images acquired by more modern (new generation) hyperspectral sensors are corrupted by noise that fits a more complicated model. In more detail, the noise has, in general, a signal-dependent (SD) nature and a SD component of the noise is prevailing compared to a signal-independent (SI) component.[14–20]

A problem is that better adequateness of the SD noise model is still often ignored in the design of methods intended for different types of hyperspectral image processing. Such processing can include denoising of junk channels,[21,22] spectral analysis,[16] and spectral feature extraction as well as compression.[19] There are some recent papers[23,24] where the authors declare their intention to take noise statistics into account when processing hyperspectral images. Meanwhile, there are also several papers where the authors have already done several steps in this direction.[16,19–22,24–26]

Obviously, to take noise statistics into account, it is necessary to have it at disposal or to estimate the statistics with appropriate accuracy. Such estimation has to be performed either from calibration data (if available) or from obtained RS data. In the latter case, it is desirable to carry out such an estimation in a blind manner. In this sense, it is worth saying that the corresponding methods are quickly developing and have been already used in several applications.[14,17,18,27–32] Moreover, the most efficient of them provides rather high accuracy.[31] Thus, a question is how to exploit this information in lossy compression of hyperspectral images?

There already exist some initial observations and results that can be employed. Skauli in Ref. 19 proposes to use standard Anscombe transform[33] and to apply it component-wise to hyperspectral images that are usually represented as 16-bit data. This simple operation practically reduces data amount twice, i.e., performs specific data compression. Moreover, such an operation makes noise in transformed sub-band images almost additive.[34] Besides, standard or generalized Anscombe transforms have been earlier exploited in lossy compression of astronomic images[35] as well as other images corrupted by SD noise.[36] This shows that such variance stabilizing transforms (VSTs) can be useful for dealing with SD noise.

Another group of observations relates to lossy compression of multidimensional data (images). Experience in color image and video compression[37] shows that if there is a correlation between multidimensional data components (e.g., frames for video), it can be used to increase the CR. Similarly, the use of three-dimensional (3-D) compression methods usually leads to sufficient increase of CR compared to component-wise compression of hyperspectral data

(see, e.g.,[38–40]). This happens due to exploiting high interchannel correlation, which is inherent for hyperspectral images.[1,4,6] However, the 3-D compression of hyperspectral data has to be carefully done taking both noise statistics and the dynamic range of sub-band images into consideration. Otherwise, decompressed images in sub-bands that have a small dynamic range can be considerably distorted.[41]

Finally, there can be different priorities of requirements to lossy compression of hyperspectral data.[4,38,40] For example, it might be required to provide CR not less than some given value with simultaneous desire to introduce less distortions into compressed data. We concentrate on another priority, the primary goal is to carry out lossy compression in the neighborhood of the optimal operation point (OOP), while it is also desirable to provide as much high CR as possible. Note that compression in the neighborhood of OOP (with efficient noise removal) practically provides maximal probability of correct classification for multichannel RS data, as this has been shown for multispectral data in Ref. 42 for the case of pure additive noise. Therefore, we can expect that similar results will be observed for the classification of hyperspectral images corrupted by SD noise.

The goal of this paper is to aggregate these observations within a unified framework or several possible approaches to lossy compression. In fact, this paper extends initial results presented in a conference paper,[25] but with distinctive differences mentioned below. Section 2 considers some peculiarities of hyperspectral data for the Hyperion sensor and presents the estimates of SD noise parameters. We give some quantitative criteria for lossy compression of single-channel noisy images that can be used in an analysis of component-wise compression of hyperspectral data. These criteria describe lossy compression from several different points of view including hyperspectral data interpreting.

Two approaches, with and without VST, to the component-wise lossy compression of images corrupted by SD noise, are presented in Sec. 3. They are tested and compared for simulated data (test images). Compared to the paper,[25] we give more details and provide a more thorough analysis on coder parameter setting in order to reach compression in an OOP neighborhood. In particular, novelty consists in analysis of different images for approaches to lossy compression. These approaches either exploit or do not use VST.

Section 4 presents the results and analysis for the case of component-wise compression applied to real-life Hyperion images. Compared to the paper,[25] more real-life data sets are considered to strengthen the conclusion drawn from an analysis of the results and to make these conclusions more general.

Section 5 discusses one approach to increase the CR by means of processing sub-band images in groups. One image is used as reference and other sub-band images are represented in different form. We present more practical examples in Ref. 25 to better stress the potentiality of this approach.

Section 6 absolutely contains new results dealing with the 3-D compression of data after VST in groups. In particular, novelty of our study consists in analysis of the influence of the group size on the efficiency of compression.

Section 7 deals with new results of analysis of endmembers for original and compressed data (we would like to thank the anonymous reviewers for their ideas to carry out such analysis). Finally, conclusions and directions of future research are given.

2 Image/Noise Model and Lossy Compression Efficiency Criteria

This paper is based on the assumption that noise in considered images is mixed when both SI and SD components are present. Thus, the noisy image can be presented as [17,27]

$$I_{ij}^n(n) = I_{ij}^{\text{true}}(n) + N_{ij}^{\text{SI}}(n) + N_{ij}^{\text{SD}}(n), \tag{1}$$

where $I_{ij}^{\text{true}}(n)$ is the true image value in the ij-th pixel of a n-th sub-band image, $N_{ij}^{\text{SI}}(n)$ and $N_{ij}^{\text{SD}}(n)$ denote SI and SD noise components in the ij-th pixel of a n-th sub-band image, $i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$, where I_{Im} and J_{Im} define the image size. These noise components are assumed zero mean, spatially uncorrelated, and Gaussian. Then, an appropriate model for the noise variance in an n-th sub-band image is $\sigma_{ij}^2(n) = \sigma_0^2(n) + k(n)I_{ij}^{\text{true}}(n)$ where $\sigma_0^2(n)$ is the SI noise variance and $k(n)$ is the SD noise parameter.

One and probably the best (the most accurate and simplest) way to estimate the parameters $\sigma_0^2(n)$ and $k(n)$ is to analyze calibration data for acquired hyperspectral images.[14] However, calibration data are not always available in practice. This means that one has to directly apply a blind method for obtaining the estimates of $\hat{\sigma}_0^2(n)$ and $\hat{k}(n)$ from acquired images. This can be done onboard if lossy compression should be performed there or on-land if lossless compression has been used before transferring the data downlink and the RS images are to be compressed with a larger CR for further transferring, dissemination, and/or storage.

Certainly, the obtained estimates of noise parameters have to be accurate enough. According to both theoretical and practical studies, the methods[14,18,43] based on the fractal Brownian motion model, maximum-likelihood estimation of noise and texture parameters in scanning windows, and noise-informative maps are able to appropriately provide the accurate estimates. Therefore, we have applied them to obtain and analyze the noise parameter estimates for the Hyperion data. In more detail, we have processed the image EO1H1800252002116110KZ[44] and some other images are available.

The obtained estimates are represented in Fig. 1 as dependences and on a sub-band index n. Since the obtained estimates of vary in wide limits ([this happens due to a variation of dynamic range and signal-to-noise ratio (SNR) in sub-band images), they are represented in a logarithmic scale. It is also worth mentioning that Hyperion data contain several sub-bands that have very narrow dynamic ranges and small SNR. Due to this, these sub-bands are commonly not used in analysis of acquired data. Because of this, we have set the estimates for such sub-bands to zero (the estimates for these sub-bands are not shown at all). Besides, these sub-band images have not been subject to compression in our further analysis.

Let us briefly analyze the dependences in Fig. 1. As seen, the estimates are within wide limits from about 10 to several thousands while are from about 0 (there are even a few estimates that are slightly negative) to about unity (mostly the estimates are of about 0.1). Both the estimates and are quite close for most neighbor sub-bands. Both of them are also quite large at the edges of sensor bands. It is worth noting here that there are two different sensors in the Hyperion RS system. The first sensor operates in visible range and the second one works in the infrared range.

Similar plots have been obtained for other Hyperion images. In particular, for the image EO1H2010262004157110KP, the same tendencies as reported above have been observed. This indirectly means that, on the one hand, parameters of the noise components in hyperspectral data for a given sensor can be quite stable. On the other hand, this means that the methods[14,18] of blind estimation are quite accurate. Note that noise has been spatially identified as uncorrelated for all analyzed data; this simplifies further simulations and analysis.

One more important observation is that using the obtained estimates, it is possible to evaluate the contributions of SD and SI noise components. Such analysis has shown that noise variance induced by an SD component is usually considerably (up to 40 times) greater than the SI noise variance at the upper margin of the data dynamic range for almost all sub-band images. Therefore, the contribution of the SD noise component is prevailing, and it should be taken into account in designing image processing methods.

For simplicity, consider a single-channel (a given sub-band) image denoted as $I_{ij}^n, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$, i.e., omitting spectral index n. Just such noisy images that
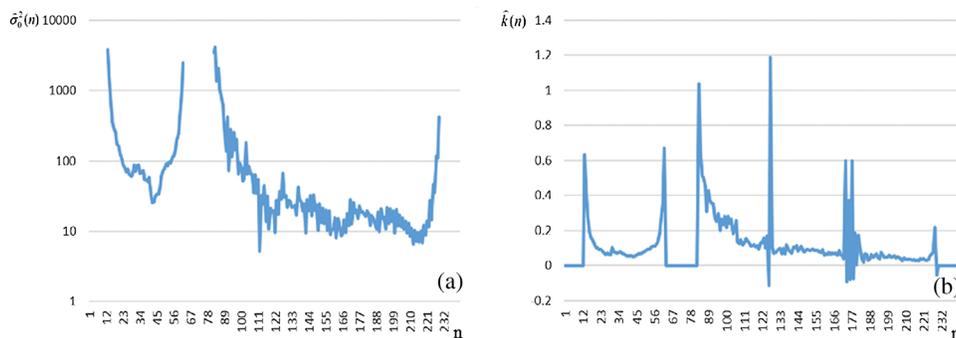
Fig. 1 Dependences (a) and (b) for the Hyperion data EO1H1800252002116110KZ.

one has available in practice. However, it is possible to have the corresponding true image also $I_{ij}^{\text{true}}, i = 1, \ldots I_{\text{Im}}; j = 1, \ldots J_{\text{Im}}$ in simulations where noisy image is obtained by generating and adding noise with preset k and σ_0^2 according to the model (1).

Suppose now that lossy compression is applied to $I_{ij}^n, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$, obtaining a compressed image $I_{ij}^c, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$. Then, in simulations, it becomes possible to calculate a metric Metr for either a pair of images I_{ij}^n and I_{ij}^c, $i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$ (denote it as Metr_{nc}) or a pair of images I_{ij}^{true} and I_{ij}^c, $i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$ (denote this metric as Metr_{tc}). Any conventional metric Metr_{nc} worsens for any coder for larger CR. This deals with the fact that more distortions are introduced in compressed data if the CR increases. Note that usually dependences of Metr_{nc} on CR [or another parameter that controls compression (PCC) as bpp or quantization step (QS) or scaling factor (SF)] are considered for conventional analysis of lossy compression performance.

Meanwhile, dependences of the metric Metr_{tc} on CR (or PCC) are specific. This has been first noticed in the paper [9] for such standard criterion (metric) as mean square error

$$\text{MSE}_{\text{tc}} = \left[\sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} (I_{ij}^c - I_{ij}^{\text{true}}) \right] / I_{\text{Im}} J_{\text{Im}}. \tag{2}$$

It has been shown that for quite intensive noise and quite simple structure images, the metric Metr_{tc} often has optimum (minimum for MSE_{tc}) for a certain value of a PCC.

Later it has been shown that other than MSE_{tc} metrics, Metr_{tc} might have optimum as well.[12] Such quality metrics, as standard peak signal-to-noise ratio (PSNR), and visual quality metrics, as multi-scale structural similarity (MSSIM)[45] and PSNR-HVS-M,[46] might have maxima. The optimum of the metric Metr_{tc} is called OOP. This optimum is associated with a PCC used for a given coder. It can be bpp_{OOP} for such coders as JPEG2000 or SPIHT, QS_{OOP} for DCT-based coders as, e.g., AGU[47] or a scaling factor SF_{OOP} for JPEG. Note that the OOP position depends upon a metric and the coder used.[12]

Below we obtain and analyze data for the coder AGU[47] that performs image compression in 32×32 pixel blocks and deblocks after decompression. There are several reasons for using this coder in our further analysis. First, it performs better than JPEG and slightly better than JPEG2000 and SPIHT.[47] Second, QS serves as PCC for this coder, and it is easy to control compression (to vary CR) by changing QS.[12,47] Third, there is also a 3-D version of AGU that has earlier been successfully applied for compressing AVIRIS hyperspectral data with sub-band grouping.[13,40]

Therefore, let us come to demonstrating and analyzing typical dependences presented in Fig. 2. These are the plots of MSE_{tc} versus QS obtained for six test images, namely, Airfield, Baboon, Barbara, Goldhill, Lenna, and Peppers. All of these test images have been corrupted by the noise with the parameters $k = 1.0$ [which is almost the largest value in
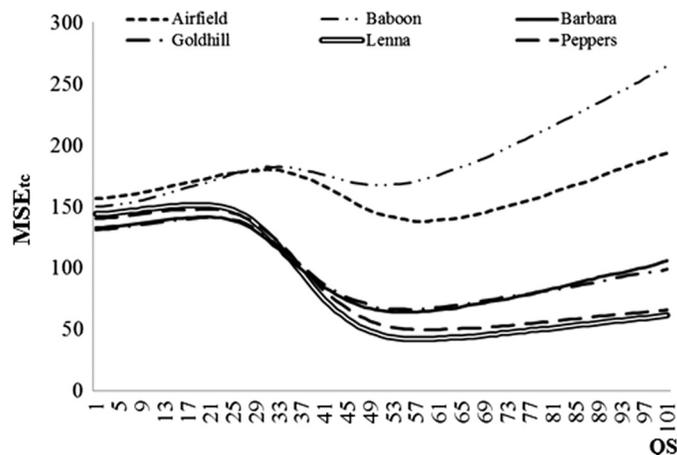
Fig. 2 Dependences $\text{MSE}_{\text{tc}}(\text{QS})$ for six test images corrupted by signal-dependent noise ($k = 1.0$ and $\sigma_0^2 = 20$).

Fig. 1(a)] and $\sigma_0^2 = 20$ [which is almost the smallest value in Fig. 1(b)]. Hence, we deal with the case of SD prevailing noise. Here, we have directly applied lossy compression to original noisy images, without any preliminary transformations.

All six dependences $\mathrm{MSE_{tc}}(\mathrm{QS})$ have global or, at least, local minima (local minimum takes place for the most textural test image, Baboon, and $\mathrm{MSE_{tc}}$ in this minimum is slightly larger than $\mathrm{MSE_{tc}}(\mathrm{QS}=1)$, i.e., for practically uncompressed images). For all these curves, minima are observed for QS about 55. Thus, we can state that OOP exists in this case, and, assuming that an image is compressed in the OOP neighborhood, the quality of I^c is closer to the quality of I^{true} than the quality of I^n (according to the analyzed metric). This brings two benefits to compression in an OOP neighborhood. The first benefit is that a compressed image has a sufficiently smaller size than the original (uncompressed or compressed in a lossless manner) image. The second benefit is that due to partial filtering of the noise, the classification of decompressed data can improve compared to classification of the original (noisy) data.[42] Moreover, classification accuracy was shown[42] to attain the optimum if multichannel data are compressed in an OOP neighborhood. These benefits give evidence in favor of compressing hyperspectral data in the neighborhood of OOP.

In practice, there can be some restrictions that might not allow carrying out the compression in an OOP neighborhood. For example, it might be necessary to ensure a desired CR and compression in OOP does not produce such a CR (produces a smaller CR than desired). In this paper, we do not analyze such a case. In fact, we do not impose any restriction on attained CR. Meanwhile, we analyze ways to provide the CR as large as possible (see Sec. 5 and 6).

The dependences in Fig. 2 have been obtained for the case of available noise-free image I^{true}. And a question is how to attain OOP (e.g., to set a proper QS for the coder AGU) if I^{true} is not available; this happens in practice. In this sense, it is reasonable to already exploit the existing experience of carrying out lossy compression in the neighborhood of OOP for other types of noise. Several procedures have been already proposed and tested for a case of additive noise.[12,48,49] The procedure in Ref. 48 is iterative and, under conditions of the known variance of additive noise, it can be used for any lossy compression method. The procedure in Refs. 12 and 49 presumes that QS is set proportional to the additive noise standard deviation according to the recommendations based on the experience obtained in advance (by simulations). Respectively, this procedure is not applicable for all coders. However, this procedure is applicable for coders for which QS or SF serve as PCC. This was one more reason for considering the AGU coder that uses QS as PCC and that has been thoroughly studied in the papers.[12,50]

It has also been shown in the papers[34,36,50] that earlier proposed automatic procedures[12,48,49] for OOP attaining can also be applied for compressing images corrupted by different types of SD noises. The main idea is in applying a proper homomorphic (variance stabilizing) image transformation before compression with converting SD noise to pure additive (or almost pure additive[34]) in transformed images. Then, under the condition that parameters of SD noise can be recalculated to a variance of additive noise, the task reduces to a simpler and already solved one.

The SD noise type determines the type of variance stabilizing transformation to be applied. For example, conventional Anscombe transform[33] performs reasonably well if noise in an original image is Poissonian. Logarithmic transform can be a proper solution if the noise in the original images is pure multiplicative.[50,51] Finally, if the noise contains both SI and SD components, the generalized Anscombe transform[35] can help in coping with the situation. Thus, taking the results of noise parameter estimation presented above into account, we pay basic attention to using the generalized Anscombe transform.[35] This is a novel approach according to which it becomes possible to perform lossy compression of images corrupted by the noise of the considered type without VST is also proposed.

3 Comparison of Approaches to Component-Wise Lossy Compression

A first approach to the compression of a one-channel (component) image corrupted by mixed noise of the model (1) can be called either direct or WithOut Variance Stabilizing Transform (WOVST). It consists in the following. Suppose that parameters k and σ_0^2 are known or pre-estimated with appropriate accuracy. Then, using k and σ_0^2 or their estimates, it is possible to determine the so-called equivalent noise variance in a given image

$$\sigma_{\text{eq}}^2 = \sigma_0^2 + \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} k I_{ij}^{\text{true}} / (I_{\text{Im}} J_{\text{Im}}) \approx \sigma_0^2 + \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} k I_{ij}^n / (I_{\text{Im}} J_{\text{Im}}). \tag{3}$$

If a coder uses QS or SF as PCC, then a coder parameter for reaching OOP can be set as

$$\text{QS} = \alpha_1 \sigma_{\text{eq}}, \text{SF} = \alpha_2 \sigma_{\text{eq}}, \tag{4}$$

where α_1 and α_2 denote some proportionality factors that depend upon the coder used. Experiments carried out by us for the coder AGU for many test images and sets of k and σ_0^2 have demonstrated that $\text{QS}_{\text{OOP}} \approx 4.5 \sigma_{\text{eq}}$. The plots in Fig. 2 confirm this. The value of σ_{eq}^2 approximately equals $\text{MSE}_{\text{tc}}(\text{QS} = 1)$, which is about 150 for all considered test images. Then, $\text{QS}_{\text{OOP}} \approx 4.5 \sigma_{\text{eq}} \approx 55$ as we have according to all plots in Fig. 2. Note that for other coders one might have other recommendations. For example, for another DCT-based coder ADCT,[52] the recommendation is to set $\text{QS} = 3.5 \sigma_{\text{eq}}$. This coder is able to produce slightly better results than AGU, but it requires more time for compressing images since the partition scheme optimization is needed.

An advantage of the WOVST approach is that compression is directly applied to the image I^n. As a result, no additional operations are needed after decompression. The corresponding lossy compression procedure can be fully automatic. At the first stage, blind estimation of noise parameters k and σ_0^2 is carried out if needed. At the second stage, σ_{eq}^2 (3) is calculated, then QS is determined according to (4) and a chosen method of lossy compression (coder) is applied with the value of QS_{OOP} set.

Now consider the second approach, the so-called WVST (With VST). According to this approach, a noisy image is first subject to the VST for obtaining the transformed image where the generalized Anscombe transform is applied (SI noise component is assumed to have zero mean in the expression below; see model (1) and comments to it)

$$I_{ij}^{\text{GA}} = (2/k) \left(k I_{ij}^n + \frac{3}{8} k^2 + \sigma_0^2 \right)^{1/2}. \tag{5}$$

If k and σ_0^2 are exactly known or accurately pre-estimated, the noise in the image I^{GA} after (5) becomes purely additive and has a variance equal to unity. This means that for coders using QS or SF as PCC one has

$$\text{QS} = \alpha_1, \text{SF} = \alpha_2, \tag{6}$$

where for the AGU coder the recommended α_1 is about 4.5.

The dynamic range of images after the transform (5) decreases and the two-dimensional (2-D) data array I_{ij}^{GA}, $i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$ becomes real-valued, although I_{ij}^n values are usually integers. This means that a coder to be applied to the data array I^{GA} should be able to deal with real-valued 2-D data. However, there is also an easy way out. If a coder is able to compress only integer-valued data, it is possible to stretch the dynamic range of I^{GA} by β times where β is adjusted to provide a desired range of data representation and to round-off the obtained values. Then, PCC can be set as

$$\text{QS} = \alpha_1 \beta, \text{SF} = \alpha_2 \beta. \tag{7}$$

In our further studies, we have not used (7) since the coders applied were able to deal with real-valued data arrays.

The decompression for the proposed approach contains more stages than the WOVST approach. After the initial decompression and deblocking (if applied), a decompressed image range is to be inversely changed (if stretching by β times has been applied before compressing). Then, the inverse generalized Anscombe transform is to be carried out.

Although it might seem difficult, the approach of WVST to image compression and decompression can be easily implemented as fully automatic. The first stage is to estimate the noise parameters k and σ_0^2 in a blind manner (if these parameters are not at disposal). Then, an image is transformed according to (5) and, if needed, stretched to a desired range. After this, QS or SF are

determined according to (6) or (7) depending upon whether its stretching is used or not. Finally, a chosen coder with preset QS or SF is applied.

The decompression is carried out in inverse order. Note that the operations of direct and inverse VST as well as data stretching/destretching are very simple and do not require essential time for their execution.

One obvious benefit of the VST-based approach to compression is clear from the description given above. For a selected coder, the used QS or SF is fixed if stretching is not used and a coder is applicable to compressing 2-D real-valued data. Note that we have modified the original coder AGU to be able to work with 2-D data.

A question now is what approach is better in the sense of providing the smallest MSE_{tc} and the largest CR in OOP? To answer it, we have obtained numerical simulation results for the considered approaches for the six above-mentioned test images corrupted by the considered type of noise. The following three sets of parameters have been analyzed: $k = 0.2$ and $\sigma_0^2 = 20$; $k = 0.4$ and $\sigma_0^2 = 20$; $k = 1$ and $\sigma_0^2 = 20$. The first set relates to comparable contributions of the SI and SD components; the other two sets correspond to the prevailing contribution of the SD component.

The obtained data are presented in Table 1. Since $\sigma_0^2 = 20$ is the same for all three sets, its values are not presented in a separate column of the table, and the values of k are given in the second column. For the approach with WOVST, the recommended QS_{rec} are given in Table 1. For the approach with preliminary VST, QS_{rec} is fixed and set equal to 4.5. The following data are presented: MSE_{tc}, $MSSIM_{tc}$, and $PSNR - HVS - M_{tc}$, all determined for the decompressed and true images. The obtained CR values are presented for all cases.

Concerning the visual quality metrics, it is worth recalling the following. The values of MSSIM vary from 0 to 1 where MSSIM = 1 relates to perfect visual quality. In turn, the values of PSNR-HVS-M are expressed in dB, and larger values of this metric correspond to a better visual quality. Note that visual quality metrics are in high correlation with the probability of correct classification for pixels that belong to small-sized objects (roads, fences, etc.).

The analysis of data presented in Table 1 shows the following:

1. If k is larger and, respectively, σ_{eq}^2 is larger, then QS_{rec} is also larger for the approach WOVST; this leads to larger values of CR provided for a given test image;
2. If k is larger, then MSE_{tc} is larger while $MSSIM_{tc}$ and PSNR-HVS-M_{tc} are smaller; these changes relate to a worse quality of decompressed images according to the considered metrics; this means that if noise intensity in the original (noisy) images is higher, then the quality of images compressed in OOP is lower; this property could be predicted in advance;
3. The value of CR reached in OOP depends upon the noise intensity and the image complexity; CR increases if noise becomes more intensive (i.e., if k and, respectively, σ_{eq}^2 are larger); similarly, CR is larger if a compressed image has a simpler structure (compare obtained values of CR for such simple structure test images as Peppers and Lenna to CR values for such textural test images as Baboon and Airfield under the condition of fixed k);
4. The tendencies abovementioned in items 2 and 3 are observed for both approaches;
5. In general, the approaches WVST and WOVST produce comparable performance (according to all considered metrics and CR) although there are small differences;
6. The CR values are slightly smaller for the WVST approach, but for this approach the values of MSE_{tc} are smaller while $MSSIM_{tc}$ and PSNR-HVS-M_{tc} are usually slightly larger; this means that the quality of images compressed in the OOP neighborhood for the WVST approach is slightly better than for the WOVST approach.

Based on the analysis performed, it is possible to conclude that, in general, both proposed approaches can be applied in practice for compressing a single-channel image or a hyperspectral image in a component-wise manner. Both WOVST and WVST approaches are able to simultaneously adapt to noise intensity and image structure. As it follows from analysis, less complex and noisier images are compressed better (with a larger CR). Meanwhile, less noisy and/or more complex structure images have to be compressed "more carefully," tending "to preserve" useful information (texture features, details, etc.) contained in such images.

Table 1 Efficiency of the considered approaches to lossy compression in optimal operation point (OOP) neighborhood for the coder AGU.

Image	k	Approach	QS_{rec}	MSE_{tc}	CR	$MSSIM_{tc}$	$PSNR - HVS - M_{tc}$
Airfield	0.2	WOVST	31.39	71.50	6.62	0.967	32.01
		WVST	4.50	69.98	6.46	0.967	32.06
	0.4	WOVST	39.57	92.57	8.16	0.955	30.02
		WVST	4.50	90.44	7.83	0.955	30.08
	1.0	WOVST	57.48	134.84	12.21	0.932	27.02
		WVST	4.50	129.75	11.37	0.933	27.19
Baboon	0.2	WOVST	30.50	76.45	5.69	0.974	32.59
		WVST	4.50	75.18	5.64	0.974	32.64
	0.4	WOVST	38.15	104.09	6.73	0.965	30.47
		WVST	4.50	102.12	6.62	0.965	30.52
	1.0	WOVST	55.06	166.05	9.19	0.944	27.17
		WVST	4.50	161.98	8.89	0.945	27.37
Barbara	0.2	WOVST	29.33	29.89	12.08	0.980	34.13
		WVST	4.50	29.85	12.06	0.980	34.19
	0.4	WOVST	36.27	38.92	14.16	0.974	32.46
		WVST	4.50	38.85	14.16	0.974	32.51
	1.0	WOVST	51.79	60.74	18.66	0.961	29.77
		WVST	4.50	59.43	18.75	0.963	29.80
Goldhill	0.2	WOVST	29.31	36.55	11.23	0.973	33.13
		WVST	4.50	35.92	11.07	0.974	33.28
	0.4	WOVST	36.25	45.94	14.01	0.964	31.33
		WVST	4.50	44.25	13.60	0.966	31.56
	1.0	WOVST	51.74	64.24	20.51	0.947	28.84
		WVST	4.50	61.21	20.01	0.951	29.06
Lenna	0.2	WOVST	30.13	22.58	18.37	0.977	34.60
		WVST	4.50	21.99	17.66	0.978	34.78
	0.4	WOVST	37.55	28.08	22.16	0.971	32.91
		WVST	4.50	27.10	20.84	0.973	33.13
	1.0	WOVST	54.02	40.10	31.00	0.960	30.45
		WVST	4.50	37.88	29.07	0.963	30.76
Peppers	0.2	WOVST	29.88	28.69	14.61	0.973	34.11
		WVST	4.50	28.22	14.01	0.973	34.17
	0.4	WOVST	37.16	34.53	18.40	0.967	32.60
		WVST	4.50	33.63	17.41	0.968	32.69
	1.0	WOVST	53.33	47.21	26.04	0.955	30.15
		WVST	4.50	46.03	24.69	0.957	30.20

Let us consider what happens in the case of lossy compression of images corrupted by SD noise. For this purpose, we have carried out an analysis for the test image airfield presented in Fig. 3(a) for the case of SD noise with $k = 1.0$ and $\sigma_0^2 = 20$. Noise is well seen in homogeneous image regions. Because of this, in our further analysis, we will pay our main attention to three quasi-homogeneous regions with different mean levels marked by digits 1, 2, and 3. Studies are carried out using difference images. The first difference image is obtained as $I_{ij}^{d1} = |I_{ij}^n - I_{ij}^{\text{true}}|, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$. It is represented in Fig. 3(b). Darker regions in this difference image correspond to areas with less intensive noise as
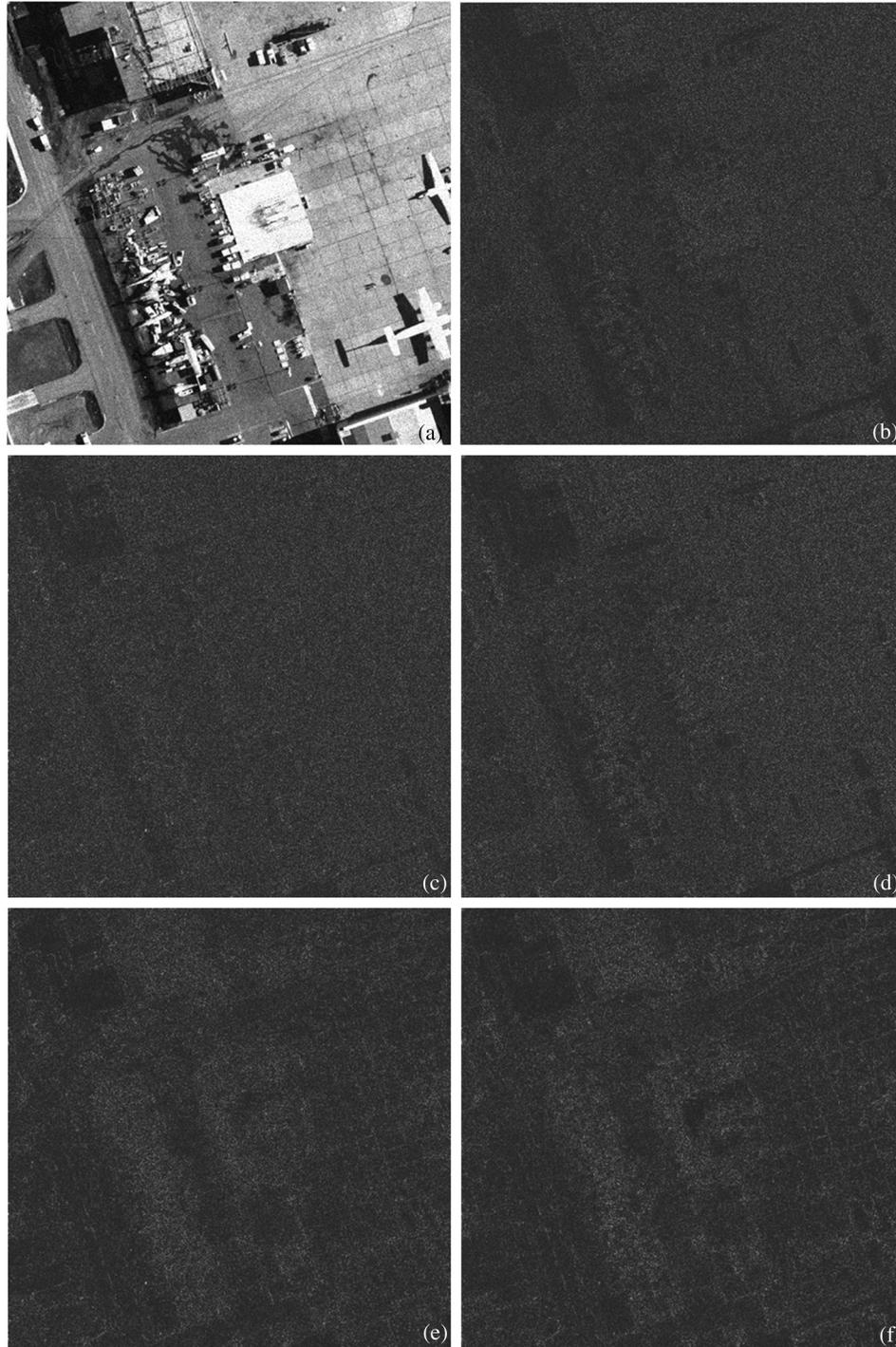
Fig. 3 Experimental data for the test image Airfield: (a) its noisy version and (b)-(f) difference images, respectively.

$I_{ij}^n, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$, i.e., areas with, on the average, smaller $I_{ij}^{\text{true}}, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$. This is well seen from a joint analysis of images in Figs. 3(a) and 3(b), in particular, in marked quasi-homogeneous regions.

We have also considered four other difference images. The second and third difference images have been obtained as $I_{ij}^{d2} = |I_{ij}^n - I_{ij}^{c\text{WOVST}}|, I_{ij}^{d2} = |I_{ij}^n - I_{ij}^{c\text{WVST}}|, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$, where $I_{ij}^{c\text{WOVST}}$ and $I_{ij}^{c\text{WVST}}$ denote the i-th pixel values of the compressed images, where VST has not been used and applied, respectively. In these images [given in Figs. 3(c) and 3(d)], one can basically analyze the properties of noise removed from noisy images by lossy compression. Compression WVST removes more noise in areas with larger mean values while lossy compression without VST introduces distortions of approximately the same level into all fragments of $I_{ij}^n, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$. These distortions deal more with noise filtering in regions with larger I_{ij}^{true} but they relate to distorting information in areas with smaller I_{ij}^{true}.

Finally, we have also obtained $I_{ij}^{d4} = |I_{ij}^{\text{true}} - I_{ij}^{c\text{WOVST}}|, I_{ij}^{d5} = |I_{ij}^{\text{true}} - I_{ij}^{c\text{WVST}}|, i = 1, \ldots, I_{\text{Im}};$ $j = 1, \ldots, J_{\text{Im}}$. These difference images characterize residual noise and introduced distortions. They are presented in Figs. 3(e) and 3(f). The images are quite similar since the values of MSE_{tc} for them are of the same order (Table 1). However, there are some differences. For fragments with larger I_{ij}^{true}, one has, on the average, I_{ij}^{d4} smaller than I_{ij}^{d5} and vice versa. This means that WVST compression approximately provides a constant factor of noise suppression in the case of SD noise. Quantitative confirmation of this can be found in the paper.[34]

We have also carried out experiments similar to those described in the section above for another coder. This coder is called ADCT[52] and has been applied within WOVST and WVST approaches. We do not present the obtained results here but give only the main observations. The tendencies described above in items 1–6 of this section are also valid for data and dependences obtained for the coder ADCT. The difference compared to AGU is the following. First, the recommended α_1 for ADCT is slightly smaller than for AGU (3.5 instead of 4.5). This slightly leads to (by 5...10%) smaller CR for ADCT compared to AGU. Second, under the abovementioned setting of QS, ADCT approximately provides the same CR as AGU for a given test image and noise parameters set. Third, MSE_{tc} values for ADCT operating in OOP are smaller by 5%–17% than the corresponding data for AGU. Similarly, MSSIM_{tc} for ADCT operating in OOP can be up to 0.08 better (larger) than for AGU, also operating in OOP. Finally, $\text{PSNR} - \text{HVS} - \text{M}_{\text{tc}}$ for ADCT can be up to 1.5 dB better than for AGU. Therefore, the quality of images compressed by ADCT in an OOP neighborhood is sufficiently better than for the corresponding images compressed by AGU. However, the coder ADCT is considerably slower, and this can restrict its application, especially for onboard compression of hyperspectral data.

4 Analysis of Component-wise Compression Efficiency for Real-Life Data

The previous two sections gave data and insight for procedures that can be used for fully automatic compression of hyperspectral images in the neighborhood of OOP. Recall that we assume noise parameters to be accurately known or pre-estimated for each sub-band. Since we know now what to do for a single-channel image, it is easy to extend the procedure to component-wise processing of multichannel (hyperspectral) data.

Consider first what to do if the WOVST approach is applied. Since the estimates of $\hat{k}(n)$ and $\hat{\sigma}_0^2(n)$ are individual (different) for each sub-band, QS or SF should also be individual. To determine PCC(n), it is necessary to use the estimates of $\hat{k}(n)$ and $\hat{\sigma}_0^2(n)$ in (3) to calculate $\sigma_{\text{eq}}^2(n)$. Then $\text{QS}_{\text{rec}}(n)$ (or $\text{SF}_{\text{rec}}(n)$) has to be calculated and remembered. The obtained set of $\text{QS}_{\text{rec}}(n), n = 1, \ldots, N$ is coded as side information to be used at the decompression stage (N denotes the number of sub-bands in hyperspectral data). Compression of each sub-band image with $\text{QS}_{\text{rec}}(n)$ (or $\text{SF}_{\text{rec}}(n)$) for a used coder is carried out afterward. Decompression is carried out in inverse order. Note that it is not necessary to remember and code $\hat{k}(n), n = 1, \ldots, N$ and $\hat{\sigma}_0^2(n), n = 1, \ldots, N$ because these arrays are not exploited at the data decompression stage.

Let us describe now the WVST approach. The estimates $\hat{k}(n)$ and $\hat{\sigma}_0^2(n)$ have to be obtained since they are used for getting the transformed images $I_{ij}^{\text{GA}}(n)$ for each sub-band. Then, lossy

compression is carried out for each sub-band image where PCC (QS or SF) is the same for all sub-bands (depending upon the coder applied and according to the recommendations known in advance). The estimate arrays of $\hat{k}(n), n = 1, \ldots, N$ and $\hat{\sigma}_0^2(n), n = 1, \ldots, N$ are to be coded as side information and added to the bit-stream. The reason for this is that these data are needed at the decompression stage to carry out an inverse transform with parameters that are individual for each sub-band.

The description given above does not take some peculiarities of real-life data into consideration. In particular, the simulated images studied in Sec. 3 have been presented as 2-D arrays of 8-bit non-negative integers (within the limits of 0,...,255). However, sub-band images of real-life hyperspectral data are usually represented by 16 bits and might also contain negative integer values. As an example, Fig. 4 gives the dependences of maximal and minimal values in the sub-band images on the sub-band index n for Hyperion data set EO1H1800252002116110KZ. Besides, some estimates $\hat{k}(n)$ can be negative [see the plot in Fig. 1(b)] although, according to theory, $k(n), n = 1, \ldots, N$ should be non-negative.

Note that similar properties of real-values data have been observed for other real-life Hyperion data sets considered by us. Thus, these peculiarities are to be taken into account in practice.

These peculiarities might cause some problems, especially for the WVST approach when executing VST. For example, it might occur that the term in (5) is negative, and this can cause problems. We propose one possible way to get around this shortcoming. Let us carry out shifting of the values $I_{ij}^n(n), i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$ as

$$I_{ij}^{ns}(n) = I_{ij}^n(n) + |I_{\min}(n)|, i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}} \tag{8}$$

for those sub-bands that have negative values of minimal value $I_{\min}(n)$ for this image.

Then, instead of conventional generalized Anscombe transform (5), the following VST form is performed

$$I_{ij}^{\text{GA}}(n) = (2/\hat{k}(n))\left(\hat{k}(n)I_{ij}^{ns}(n) + \frac{3}{8}\hat{k}^2(n) + \hat{\sigma}_0^2(n) + \hat{k}(n)|I_{\min}(n)|\right)^{1/2}, i = 1, \ldots, I_{\text{Im}}; j$$

$$= 1, \ldots, J_{\text{Im}}. \tag{9}$$

The values $I_{\min}(n)$ have to be remembered and coded as ancillary information alongside the sets of $\hat{k}(n)$ and $\hat{\sigma}_0^2(n)$.

Besides, if, for a given sub-band, one has $\hat{k}(n) < 0$, then $\hat{k}(n) = 0$ is assumed and another modification is used, namely $I_{ij}^{\text{GA}}(n) = I_{ij}^n(n)/\hat{\sigma}_0(n), i = 1, \ldots, I_{\text{Im}}; j = 1, \ldots, J_{\text{Im}}$. This means that image value shifting is not carried out, but data normalization is performed instead to obtain the variance of noise close to unity. The introduced modifications can be easily and quickly done at the initial stage of hyperspectral data analysis to prepare them to compression.

Data decompression contains the following stages. First, ancillary information (the sets of $\hat{k}(n)$, $\hat{\sigma}_0^2(n)$, and $I_{\min}(n)$, $n = 1, \ldots, N$) are decoded. Then, for sub-bands having $\hat{k}(n) < 0$,
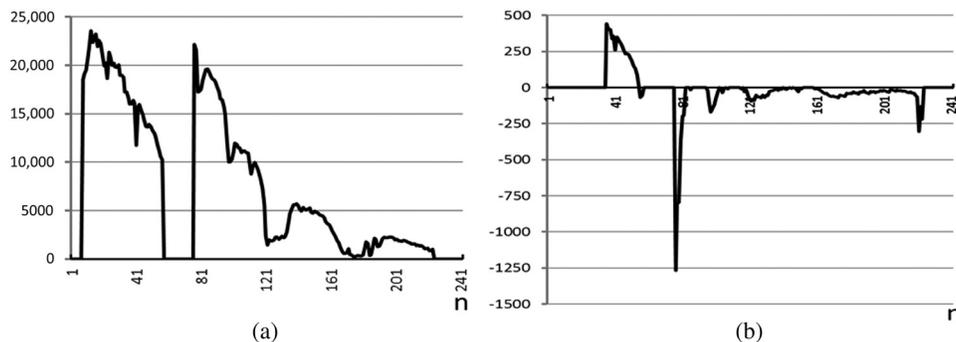
Fig. 4 (a) Maximal and (b) minimal values versus sub-band index n for sub-band images, data set EO1H1800252002116110KZ.

the final decompressed images are obtained as $I_{ij}^d(n) = I_{ij}^{\text{GA}d}(n)\hat{\sigma}_0(n)$, $i = 1,\ldots,I_{\text{Im}}$; $j = 1,\ldots,J_{\text{Im}}$. Here $\{I_{ij}^{\text{GA}d}(n)\}$ denotes the originally decompressed image for the n-th sub-band. This means that decompressed data are simply stretched to the original range.

For the sub-bands having negative $I_{\min}(n)$, inverse transform and range shifting are performed according to the following expression:

$$I_{ij}^d(n) = (I_{ij}^{\text{GA}d}(n)/2)^2\hat{k}(n) - \frac{3}{8}\hat{k}(n) - \hat{\sigma}_0^2(n)/\hat{k}(n) - |I_{\min}(n)|, i = 1,\ldots,I_{\text{Im}}; j = 1,\ldots,J_{\text{Im}}.$$

(10)

Finally, for sub-bands with non-negative $\hat{k}(n)$ and $I_{\min}(n)$, the final decompressed data are obtained as follows:

$$I_{ij}^d(n) = (I_{ij}^{\text{GA}d}(n)/2)^2\hat{k}(n) - \frac{3}{8}\hat{k}(n) - \hat{\sigma}_0^2(n)/\hat{k}(n), i = 1,\ldots,I_{\text{Im}}; j = 1,\ldots,J_{\text{Im}},$$

(11)

i.e., using inverse VST in its original form that corresponds to direct VST (5).

We have analyzed both WOVST and WVST component-wise lossy compression of real-life hyperspectral data. Both approaches have led to comparable (very similar) results. Because of this, below we present data only for the approach of WVST.

This has been done[25] for the Hyperion hyperspectral data EO1H1800252002116110KZ. Recall that for real-life data, the true sub-band images are not available. Hence, compression performance can be described by the attained $\text{CR}(n)$ as well as by losses introduced by compression. These losses can be characterized by $\text{MSE}_{\text{nc}}(n)$ calculated for the original $\{I_{ij}^n(n)\}$ and compressed $\{I_{ij}^c(n)\}$ images in each sub-band.

Although the recommended QS for the coder AGU used within WVST approach is equal to 4.5, we have also used some other values of QS. The obtained plots of $\text{MSE}_{\text{nc}}(n)$ are represented in Fig. 5 in logarithmic scale, three values of QS, namely, 2.5, 3.5, and 4.5 are considered. Besides, we also present the plot of $\sigma_{\text{eq}}^2(n)$.

As expected, an increase of QS leads to larger $\text{MSE}_{\text{nc}}(n)$ for any given sub-band. However, values of $\text{MSE}_{\text{nc}}(n)$ for losses introduced in different sub-bands differ considerably, up to three orders. And $\text{MSE}_{\text{nc}}(n)$ are of the same order as $\sigma_{\text{eq}}^2(n)$ for each sub-band. Moreover, for QS = 3.5, the plots of $\text{MSE}_{\text{nc}}(n)$ and $\sigma_{\text{eq}}^2(n)$ practically coincide.

These are interesting and important observations but let us further check what happens for other real-life data sets. Figure 6 presents the plots of $\text{MSE}_{\text{nc}}(n)$ and $\sigma_{\text{eq}}^2(n)$ for another real-life Hyperion data set, only the case QS = 3.5 is analyzed. As seen, the shapes of $\sigma_{\text{eq}}^2(n)$ dependences in Figs. 5 and 6 are very similar. And again the plots $\text{MSE}_{\text{nc}}(n)$ and $\sigma_{\text{eq}}^2(n)$ for QS = 3.5 practically coincide.

This is not surprising. First, this means that the coder introduces the losses that mainly relate to noise filtering. Second, similar effects were earlier observed for lossy compression of images corrupted by additive noise:[48] MSE_{nc} in the neighborhood of OOP and additive noise variance were approximately equal. Moreover, this property was put into the basis of automatic iterative
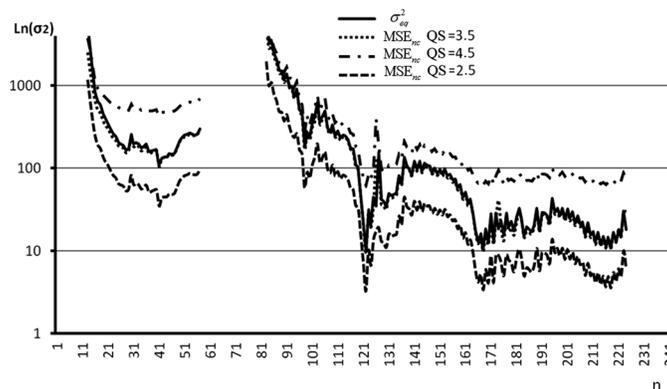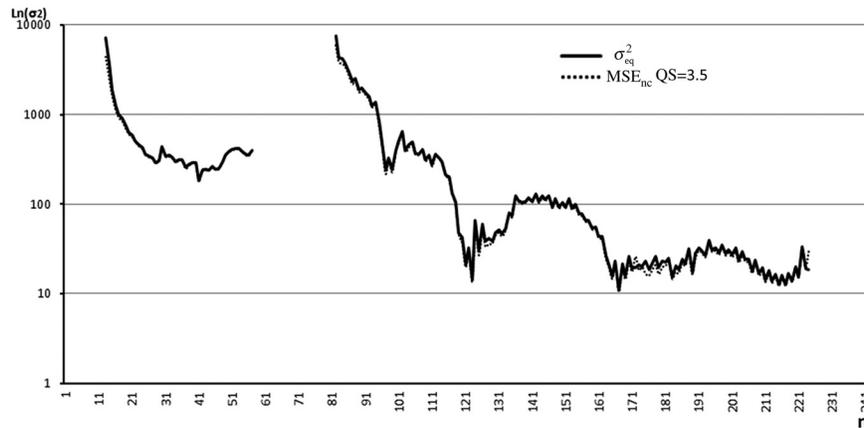

Fig. 5 Dependences (for three values of QS) and data set EO1H1800252002116110KZ.

Fig. 6 Dependences (QS = 3.5) and data set EO1H2010262004157110KP.

procedure of OOP attained in Ref. 48. Thus, the approximate coincidence of the plots $\mathrm{MSE}_{nc}(n)$ and $\sigma_{eq}^2(n)$ indirectly confirms that we have compressed data in the neighborhood of OOP.

The earlier simulation results and compression data for real-life images show that QS for AGU in practice should be about 4. A slightly smaller value (QS = 3.5) is recommended for "more careful" lossy compressions when the primary task is not to introduce too many losses. In turn, slightly larger QS (e.g., QS = 4.5) can be used if it is desirable to provide a larger CR. In fact, QS plays the role of PCC for the considered case of component-wise lossy compression of hyperspectral data.

Now consider the dependences of CR(n) for the proposed component-wise automatic compression procedure WVST based on an AGU coder. The data are presented for lossy compression with QS = 3.5 (see Fig. 7, solid lines, denoted as AGU with VST). For comparison purposes, we also give data for the lossless coder (archiver) RAR applied component-wise (dash-dot lines, denoted as RAR). Figure 7 represents the plots for two considered real-life data sets.

As it follows from the analysis, CR(n) for the proposed procedure varies from about 4.5 to 25 [Fig. 7(a)] and from 4.5 to 40 [Fig. 7(b)]. CR values are about 6.0 for most sub-band images of the first data set and about 7.0 for the second data set. In turn, the values of CR(n) for archiver RAR are considerably smaller, and they vary from about 1.5 to 2.0 for both data sets. Therefore, the proposed procedure provides improvement of CR by, at least, 3 times compared to the lossless counterpart.

It could be interesting to look at what kind of losses are introduced into compressed images. Processed hyperspectral images are of quite a large size $I_{Im} \times J_{Im} = 256 \times 3072$ pixels. Therefore, we have cut from them only some of the most informative fragments. Figure 8 presents the original and compressed images for the 220'th sub-band of hyperspectral image (data set EO1H1800252002116110KZ). As seen, noise is noticeable in the original image [Fig. 8(a)].

In the compressed image [Fig. 8(b)], valuable information at the edges of small-sized objects and textures, are not distorted while the noise is partly suppressed. The effects of noise filtering are well seen in homogeneous image regions. Therefore, the desired positive effects are attained for this component image compressed with CR about 6.0; see Fig. 7(a).

Another example is shown in Fig. 9. This is the sub-band that can be considered good (high quality), i.e., characterized by high SNR and practical invisibility of the noise in the original data [see Fig. 9(a)]. Concerning the compressed image [see Fig. 9(b)], the introduced distortions are not observed by visual inspection, and both images in Fig. 9 look identical. CR for this sub-band image is about 5.5, i.e., quite large. In fact, for this sub-band, one deals visually with lossless compression described in several papers dealing with lossy compression of medical and optical images. [53–55]

The presented examples demonstrate that the proposed approach to component-wise lossy compression performs well enough, providing reasonable filtering of images in components with relatively small SNR and "invisible" distortions into sub-band images with relatively large SNR. The question now is whether or not CR can be further increased without increasing the distortion level.
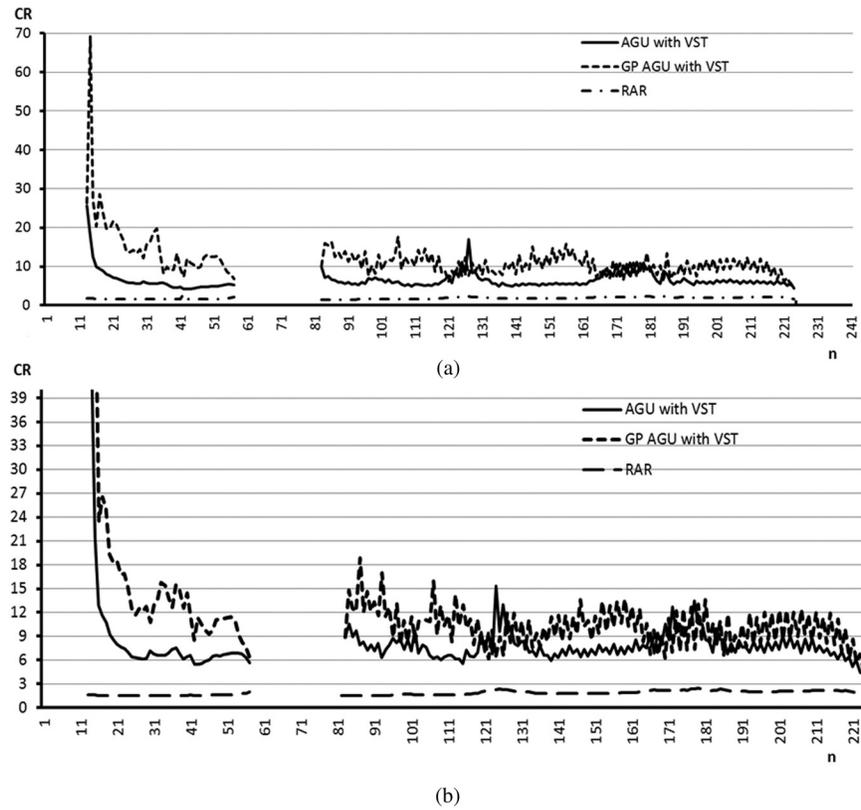
Fig. 7 Dependences on sub-band index for two proposed procedures and RAR:(a) data set EO1H1800252002116110KZ and (b) data set EO1H2010262004157110KP.

5 Compression of Difference Images with Sub-Band Grouping

The component-wise compression procedure described above does not exploit the interband correlation of hyperspectral data known to be high.[1,4,5] Meanwhile, this inherent property allows increasing CR by employing spectral redundancy of the data without increasing losses.[4,5,13,38–40] There are numerous possible approaches to carry out 3-D compression. Our goal here is not to compare them and/or to find the best (optimal) approach. Instead, we would like to test how
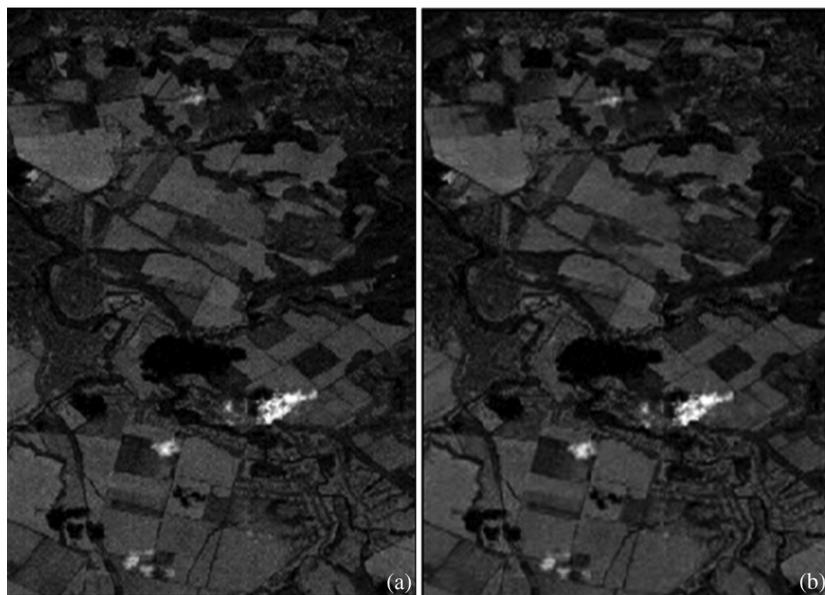
Fig. 8 Fragments of (a) original and (b) compressed image in 220'th sub-band.
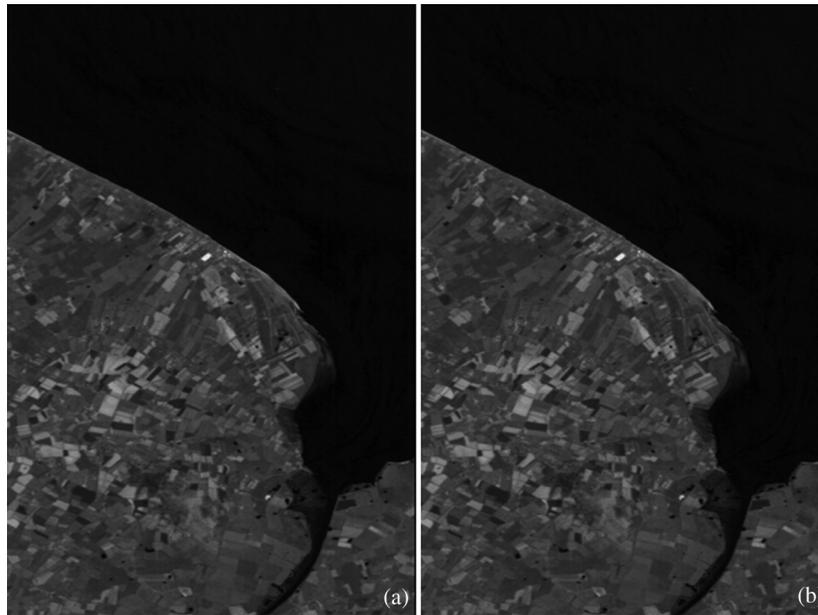
Fig. 9 Fragments of (a) original and (b) compressed image in 145'th sub-band.

efficient the 3-D compression could be in the case of VST-based image preprocessing. Besides, we would like to show what benefits can be provided by such preprocessing.

It has already been demonstrated in the paper[41] that setting QS or SF that are too large for sub-band images with small dynamic range. When these images are included into a 3-D group and jointly compressed, this can lead to destroying data in such sub-bands after decompression (due to lossy compression).

In this sense, VST applied component-wise before compression provides several favorable prerequisites. First, noise becomes additive and has practically the same intensity in all sub-bands (some differences in variance of the noise in sub-band images can be detected due to errors of SD noise parameters estimation). Second, dynamic ranges of sub-band images after direct VST become closer, and this allows avoiding the abovementioned drawback of 3-D compression.

The abovementioned prerequisites can be exploited in several ways dealing with setting fixed QS or SF for groups of sub-band images. Note that there are many ways for forming such groups. For example, for Hyperion data there can be two groups where the first one includes sub-band images from one (optical) sensor, and the second group contains sub-band data for an infrared sensor. It is also possible to have more groups containing fixed or individual numbers of sub-bands for each group. Finally, all sub-band images can be considered as 3-D data array and compressed together.

Having this in mind, we have decided not to analyze optimality of data groupings here. In this section, we consider one way of lossy compression while the next section deals with another way. And even for the second way, which provides better results, we cannot guarantee that it is the best. Our primary goal is only to demonstrate the sufficient increase in CR due to exploiting interband correlations.

Thus, let us assume that the first stage of the proposed procedure is the same as earlier, i.e., for component-wise preprocessing. After determining the parameters of the noise in all sub-bands, the corresponding VST, image shifting, and/or normalization (if needed) are carried out (see the corresponding expressions in the previous section). Then, we propose to divide this hyperspectral image into a few groups. Below we examine the case of two groups where the first one contains sub-bands with indices from 13 to 57, and the second group includes sub-band images with indices from 83 to 224. This means that we process and further compress only essential Hyperion data (see Fig. 1 and the discussion at the beginning of Sec. 2).

For each group, we consider the first sub-band image as a reference. For other images in each group, the following operations with obtained difference images are performed:

1. The first (reference) image $I^{GA}(1)$ is compressed with a set QS (recommended QS = 3.5).

2. For the second image in the considered group, the difference image is obtained as $I_{ij}^{ns\delta}(2) = I_{ij}^{GA}(1) - I_{ij}^{GA}(2)$, $i = 1, \ldots, I_{Im}; j = 1, \ldots, J_{Im}$; it is then compressed using QS = 4.0.

3. The third and other images in the group are obtained as follows:

   a. earlier compressed difference images $I^{ns\delta}(q-1)$ are decompressed by obtaining $I^{ns\delta d}(q-1)$. q denotes the sub-band image index in a group and Q defines the number of sub-band images in this group;

   b. the supplementary image is calculated as $I_{ij}^{dec}(q-1) = I_{ij}^{GA}(q-2) - I_{ij}^{ns\delta d}(q-1)$, $i = 1, \ldots, I_{Im}; j = 1, \ldots, J_{Im}$;

   c. the difference image is derived as $I_{ij}^{ns\delta}(q) = I_{ij}^{dec}(q-1) - I_{ij}^{GA}(q), , q = 3, \ldots, Q.$

   d. the obtained difference image $I^{ns\delta}(q)$ is compressed using QS = 4.0.

Due to the use of decompressed images in the calculation of difference, there is no accumulation of errors.

Let us explain the main idea of this approach. Due to the considerable interband correlation and applied VST, the difference images have smaller dynamic ranges than the corresponding reference image. If the interband correlation is very high, then difference images mostly have a noise-like appearance (similar effects take place in frames of video when there is practically no motion between neighbor frames[37]). Small dynamic ranges of the data and their noise-like structure allow their efficient compression because many DCT coefficients in blocks after quantization become quite equal to zero and such data are well coded. Note that if noise in neighbor images $I^{GA}(q)$ and $I^{GA}(q-1)$ is uncorrelated, then the noise variance in the difference image is larger than in each of these sub-band images. Then, according to the results of studies in Sec. 2 and papers,[48,49] QS for the difference images should be larger than for the reference images.

Let us partly illustrate the aforesaid by images and their characteristics. Consider a fragment of the 20'th and 19'th sub-bands for the first Hyperion subset. In the original images (before VST), minimal values for both sub-bands are equal to zero and maximal values are on the order of 17,000. After VST in these sub-bands, minimal values became equal at 269 and 243 while maximal became 875 and 804 for the 20'th and 19'th sub-bands, respectively. Thus, the dynamic range has considerably changed to a narrower one after VST. The 20'th sub-band image after VST and image normalization (dynamic range compressing to fit the interval 0–255) is shown in Fig. 10(a). The difference image has the range from 25 to 72, i.e., it is considerably smaller than for images after VST. This difference image, without any changes of dynamic range, is presented in Fig. 10(b). It really has many quasi-homogeneous regions with little noise (its variance is about 2). Meanwhile, this difference image contains "some objects" that have much less contrast with respect to the surrounding background than the corresponding contrasts in the sub-band images (after VST).

Decompression is performed in inverse order. First, the reference and difference images are decompressed. Then, decompressed sub-band images are calculated for nonreference sub-bands as $I_{ij}^{dec}(q) = I_{ij}^{GAd}(q-1) - I_{ij}^{ns\delta d}(q), i = 1, \ldots, I_{Im}; j = 1, \ldots, J_{Im}$ where I^{GAd} defines the decoded sub-band image. After this, the inverse transform is carried out individually for each sub-band image according to the rules described above in Sec. 4. As seen, both procedures of compression and decompression can be fully automatic.

The dependence $MSE_{nc}(n)$ has been obtained for the considered approach denoted as group processing (GP) AGU with VST. These dependences are practically the same as those for component-wise compression with QS = 3.5 [see the corresponding plots in Fig. (6)]. Hence, it is possible to conclude that compression is provided in the neighborhood of OOP for each sub-band.

Consider now dependences $CR(n)$. They have already been presented above by dashed lines in Fig. 7. As seen, the compression procedure GP AGU WVST results in increasing the CR for almost all sub-bands where difference images are compressed instead of $\{I_{ij}^{GA}(n)\}$. The total CR for hyperspectral data increases by about 2 times with respect to the component-wise compression. This happens for both data sets presented in Fig. 7, and the same improvement has been
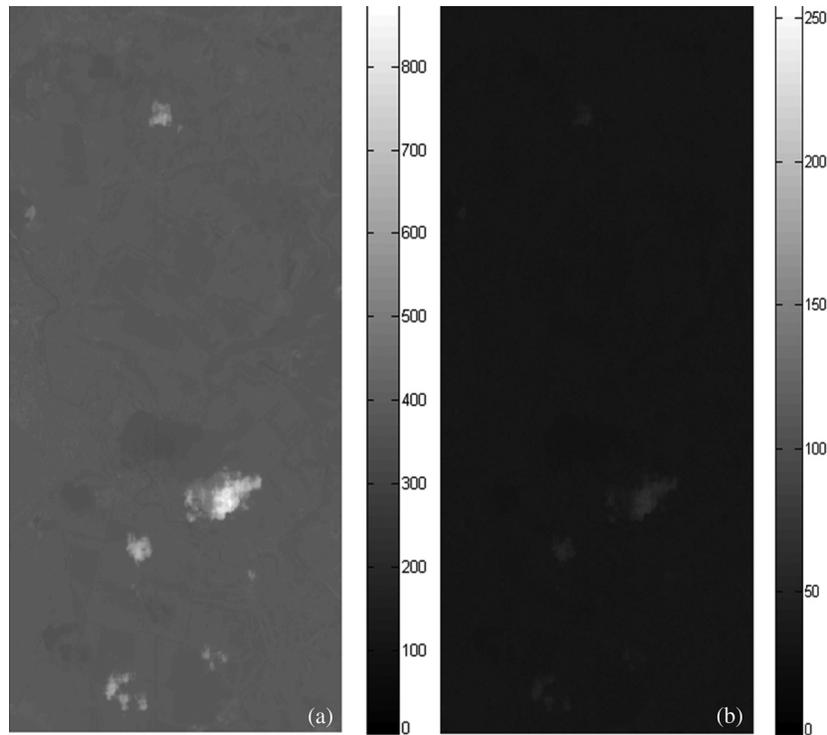
Fig. 10 Sub-band image (a) after VST and (b) the difference image.

observed for other Hyperion data sets compressed by us. Therefore, we can state that by exploiting the interchannel correlation of the data, one is sufficiently able to increase CR without increasing the level of losses introduced by lossy compression.

6 Three-Dimensional Compression with Sub-Band Grouping

The approach proposed and described in Sec. 5 exploits interchannel correlation but, in fact, it does not carry out the actual 3-D compression by processing hyperspectral images as 3-D data arrays. Below we consider three ways of such compression, all applied to images after VST described in Sec. 4:

Sub-band images are collected in groups of size 8 sub-bands (denoted as AGU8WVST);

Sub-band data are combined in 16 sub-band group (AGU16WVST);

All sub-band images are represented as a joint 3-D data array and then compressed (AGUAllWVST).

Note that we have applied all these approaches separately for optical and infrared sensors data subsets, the first contains 43 sub-bands and the other has 142 sub-bands. Since there are remaining sub-bands for both subsets for groups of size 8 and 16, these sub-bands have been collected into smaller size groups.

Compression has been done using 3-D modification of the AGU coder earlier employed and studied in Refs. 13 and 40 with application to AVIRIS data. The 3-D AGU carries out the DCT for spectral decorrelation of the data in each group and then performs as for the 2-D case. In Ref. 13, groups of sub-band images have been adaptively formed based on the analysis of noise variance in neighbor sub-bands. In Ref. 40, sub-band images have been normalized before compression using the estimates of noise variance. In both cases, the noise was supposed pure additive.

In our case, noise is also assumed additive, but it becomes such after applying VST with properly adjusted parameters. For 3-D AGU, recommendations concerning the setting QS are the same as for the 2-D case. Because of this, we used QS = 3.5 in all our experiments described in this section. To demonstrate that this is a proper choice, Fig. 11 presents $\mathrm{MSE_{nc}}(n)$ and $\sigma_{\mathrm{eq}}^2(n)$ dependences for the approach AGU8WVST on the first Hyperion data

Fig. 11 Dependences (QS = 3.5, approach AGU8WVST) and, data set EO1H1800252002116110KZ.

set. These dependences practically coincide, and this shows that compression is performed in the neighborhood of OOP for all sub-band images.

Dependences $MSE_{nc}(n)$ have also been obtained for the approaches AGU16WVST and AGUAllWVST. They are very similar to the one represented in Fig. 11 for the approach AGU8WVST. Similarly, we also obtained dependences $MSE_{nc}(n)$ for the second considered Hyperion data set. These dependences have been compared between each other as well as to σ_{eq}^2. All four dependences are very similar. Thus, we deal with compression in the OOP neighborhood and can compare the analyzed approaches from the viewpoint of the provided CR.

One way to present dependences of CR is used in Fig. 12. Note that CR cannot be individually determined for each sub-band since sub-band images are compressed in groups. Thus, we
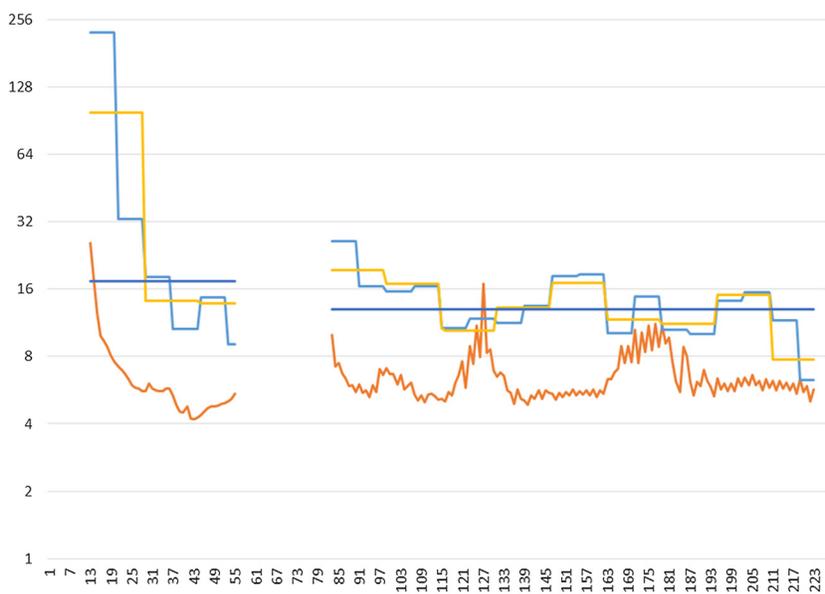
Fig. 12 Dependences CR(n) for approaches AGU8WVST (blue), AGU16WVST (yellow), AGUAllWVST (violet) and component WVST (red) for the data set EO1H1800252002116110KZ.

Table 2 Values of CR for different three-dimensional (3-D) approaches to lossy compression for two considered data sets.

Approaches to compression	Data set EO1H1800252002116110KZ		Data set EO1H2010262004157110KP	
	Optical	Infrared	Optical	Infrared
AGU8WVST	18.34	12.82	18.89	14.26
AGU16WVST	20.81	14.65	25.62	14.02
AGUAllWVST	17.43	13.00	34.97	13.40

assign CR attained for a given group to all sub-band images of this group and show them by horizontal segments. For comparison purposes, CR(n) for component-wise compression with VST is also presented. As seen, the values of CR vary from one group to another. Larger CRs are usually provided for groups of sub-bands having smaller SNR as for the group starting from $n = 13$.

The obtained values of CR have been collected in Table 2 to compare performances for the considered approaches. Data analysis shows the following. Optical subset images are compressed better than infrared images. Compression with groups containing 16 sub-bands is more efficient than for the approaches AGU8WVST and AGUAllWVST. Note that similar effects have been earlier observed in the paper.[40] It is quite difficult to explain this at the moment. Probably, the reason is in the correlation degree between neighbor sub-bands that decrease sub-bands with more distant wavelengths (indices).

In any case, 3-D compression considerably provides (several times) larger CR than in the case of component-wise compression.

7 End-Member Analysis

In this section, we will show that lossy compression does not introduce negative effects in the unmixing task, under the hypothesis that the linear setting can be admitted in practice. Recall that the linear mixing model (LMM) assumes that the radiance or reflectance spectrum associated with each pixel can be locally viewed as a linear combination of a limited number of distinct materials, commonly known as endmembers.[56] The fractions, in which these materials appear in an observed mixed pixel, are called fractional abundances. The LMM hypothesis especially holds when the macroscopic mixture of distinct materials, with relatively constant spectral properties, takes place for each observed pixel in the area of interest. In practice, when endmembers are unknown, end-member signatures should be extracted or estimated first. However, some end-member extraction algorithms can simultaneously provide abundance estimates.

Consider the end-member analysis carried out on the previously considered Hyperion data set using two different and popular methods for extracting endmembers among observed data such as the vertex component analysis (VCA)[57] and the sequential maximum angle convex cone (SMACC) method.[56] Both methods are sequential, producing a set of endmembers in sequential order. Once the number of endmembers and their spectral signatures are initially extracted by virtual dimensionality (VD) and VCA, abundances can be subsequently estimated via the least-squares approach with or without additional constraints (abundance non-negativity and/or sum-to-one constraints). SMACC provides abundance images simultaneously to endmembers to determine the fractions of the total spectrally integrated radiance or reflectance of a pixel contributed by each resulting endmember. Whether abundance constraints should be imposed or not depends on practical application. It has been argued that if the number of endmembers and their signatures are accurate, the two constraints should be automatically satisfied.[58] Note that the non-negativity constraint is more important than the sum-to-one constraint. Due to noise and spectral variability, reinforcing the sum-to-one constraint may be prone to induce additional estimation error.

We have performed our analysis in the following manner. We have considered the three cases investigated above in the paper: the Hyperion original image and its two lossy compressed versions (WOVST and WVST). These data sets have first been cropped to obtain reduced size
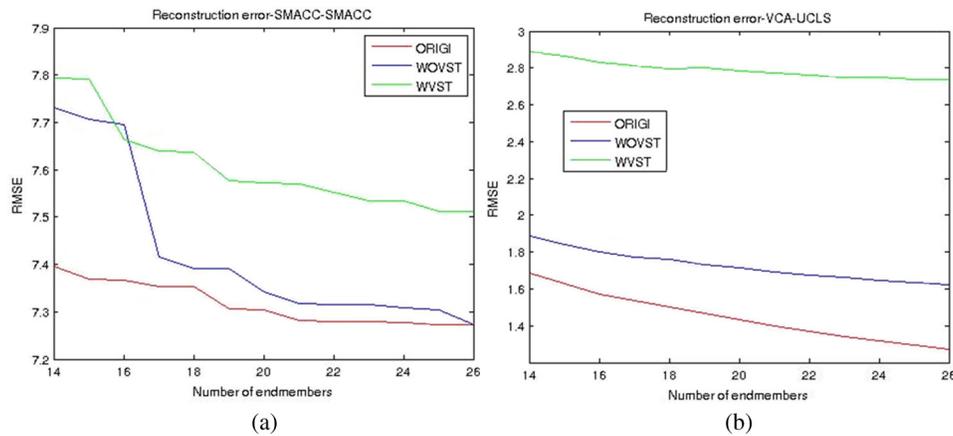
Fig. 13 RMSE versus (a) endmembers number for VCA and UCLS or (b) SMACC approaches for three considered sub-images, Original, WOVST, and VST lossy compressed.

sub-images (each with a size of $800 \times 256 \times 187$ with spectral indices from 13 to 57 and from 83 to 224) containing a minimum number, i.e., two, of very small clouds. Abundance maps have been derived both from the unconstrained least-squares (UCLS) linear unmixing method and the SMACC method for an increasing number of endmembers within a specific interval (from 14 to 26), binding the estimated virtual dimensionality of the lossy compressed sub-images.

Unfortunately, the evaluation of unmixing is a difficult task, and this is also confirmed here due to the unavailability of ground-truth data. Thus, we have decided to resort to the standard pixel reconstruction error for comparing the linear decomposition obtained on the original and compressed sub-images. However, we have to stress that the pixel reconstruction error obtained after direct unmixing of the original noisy sub-image is just a rough reference, not totally informative for our purpose, as it mainly translates the error introduced by the linear unmixing process itself while being applied on a noisy observation.

The obtained results are presented in Fig. 13. Each figure shows one specific approach, either VCA and UCLS or SMACC, the dependencies of the mean pixel reconstruction error according to the number of endmembers considered. The average is calculated based on the whole spatial extent available. The general conclusions we can give are the following. Pixel reconstruction error decreases when using more endmembers for both approaches, as is generally the case for all unmixing and especially linear methods. The level of the mean pixel reconstruction error after lossy compression remains close to that obtained without lossy compression, meaning that the lossy compression approaches operating at OOP are consistent in terms of preserving useful information and mainly removing noise.

Note that we have not considered methods imposing more constraints in the unmixing process (mentioned above), as the pixel reconstruction error is increased in such a case, at a level larger than the one observed in the difference between the pixel reconstruction error for the original and lossy compressed sub-images. Concerning the SMACC method, the pixel reconstruction errors for the original sub-image and the lossy compressed one without VST are quite similar for a number of endmembers larger than 17. The evaluation could be more reasonable in terms of abundance accuracy. But it cannot be performed here due to the unavailability of ground-truth data.

The first results obtained here need to be extended in the future based on the experiments on synthetic data to better discriminate in the pixel reconstruction error and in the accuracy of abundance maps the relative influence of unmixing and lossy compression, both in linear and nonlinear unmixing frameworks.

8 Conclusions and Future Work

Several approaches to component-wise and 3-D lossy compression of hyperspectral data are proposed and studied. The novelty is that all of them in one way or another take into

consideration the prevailing contribution of the SD noise component in sub-band images. The performance of these approaches is studied and compared.

In the case of component-wise compression, two approaches are proposed and investigated. One of them exploits VST while another approach does not. It is shown that both approaches provide comparable performance from the viewpoint of decompressed image quality and CR reached for data compressing in the neighborhood of optimal operation point. Fully automatic procedures for such compressions are described.

A set of original 3-D lossy compression approaches are considered. All of them exploit inter-channel correlation in hyperspectral data, and all employ VST as a useful preprocessing operation. The use of VST allows easy reaching of OOP and compression parameter settings. All of these approaches provide considerable benefit in CR with respect to component-wise compression under conditions of the same level of introduced distortions. Initial studies concerning the size of the groups are carried out. It is demonstrated that the group size equal to 16 could be a good practical choice. The fact that the proposed approaches are fully automatic makes them useful for applying onboard a hyperspectral sensor carrier, both airborne and space-borne.

End-member analysis for original and compressed data is performed as well. It shows that the proposed approach to compression does not introduce distortions that can be crucial for solving the final tasks of hyperspectral data processing as, e.g., unmixing.

Acknowledgments

References

1. C.-I. Chang, Ed., *Hyperspectral Data Exploitation: Theory and Applications*, 1st ed., pp. 440, Wiley-Interscience, Hoboken, New Jersey (2007).
2. R. A. Schowengerdt, Ed., *Remote Sensing Models and Methods for Image Processing*, 3rd ed., pp. 530, Academic Press, Inc., Orlando, Florida (2007).
3. A. Plaza et al., "Foreword to the special issue on hyperspectral image and signal processing," *IEEE J. Sel. Topics Appl. Earth Observ.* **5**(2), 347–353 (2012).
4. E. Christophe, "Hyperspectral data compression tradeoff," in *Optical Remote Sensing, Advances in Signal Processing and Exploitation Techniques*, S. Prasad, L. M. Bruce, and J. Chanussot, Eds., pp. 9–29, Springer, Berlin, Germany (2011).
5. A. Kaarna, "Compression of Spectral Images," Chapter 14 in *Vision Systems: Segmentation and Pattern Recognition*, G. Ohinata and A. Dutta, Eds., pp. 269–298, I-Tech, Vienna, Austria (2007).
6. E. Magli, "Multiband lossless compression of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.* **47**(4), 1168–1178 (2009).
7. M. J. Ryan and J. F. Arnold, "Lossy compression of hyperspectral data using vector quantization," *Remote Sens. Environ.* **61**(3), 419–436 (1997).
8. L. Bo, "Remote-sensing image compression using two-dimensional oriented wavelet transform," *IEEE Trans. Geosci. Remote Sens.* **49**(1), 236–250 (2011).
9. O. K. Al-Chaykh and R. M. Mersereau, "Lossy compression of noisy images," *IEEE Trans. Image Process.* **7**(12), 1641–1652 (1998).
10. B. Aiazzi, L. Alparone, and S. Baronti, "Near-lossless compression of 3-D optical data," *IEEE Trans. Geosci. Remote Sens.* **39**(11), 2547–2557 (2001).
11. I. Blanes et al., "Classification of hyperspectral images compressed through 3D-JPEG2000," in *Proc. LNAI*, Vol. 5179, pp. 416–423, Springer, Berlin, Heidelberg (2009).
12. N. Ponomarenko et al., "Lossy compression of noisy images based on visual quality: a comprehensive study," *EURASIP J. Advan. Signal Process.* **2010**, 976436 (2010).
13. V. Lukin et al., "An automatic approach to lossy compression of AVIRIS hyperspectral data," *Telecommun. Radio Eng.* **69**(6), 537–563 (2010).

14. M. Uss et al., "Local signal-dependent noise variance estimation from hyperspectral textural images," *IEEE J. Sel. Topics Signal Process.* **5**(3), 469–486 (2011).
15. A. Barducci et al., "CHRIS-Proba performance evaluation: signal-to-noise ratio, instrument efficiency and data quality from acquisitions over San Rossore (Italy) test site," in *Proc. 3rd ESA CHRIS/Proba Workshop*, ESRIN, Frascati, Italy (2005).
16. B. Aiazzi et al., "Benefits of signal-dependent noise reduction for spectral analysis of data from advanced imaging spectrometers," in *Proc. 3rd Workshop on Hyperspectral Signal and Image Processing: Evolution in Remote Sensing*, pp. 1–4 (2011).
17. J. Meola et al., "Modeling and estimation of signal-dependent noise in hyperspectral imagery," *Appl. Opt.* **50**(21), 3829–3846 (2011).
18. M. L. Uss et al., "Maximum likelihood estimation of spatially correlated signal-dependent noise in hyperspectral images," *Opt. Eng.* **51**(11), 111712 (2012).
19. T. Skauli, "Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing," *Opt. Express* **19**(14), 13031–13046 (2011).
20. V. Lukin et al., "Methods and automatic procedures for processing images based on blind evaluation of noise type and characteristics," *J. Appl. Remote Sens.* **5**(1), 053502 (2011).
21. P. Zhong and R. Wang, "Multiple-spectral-band CRFs for denoising junk bands of hyper-spectral imagery," *IEEE Trans. Geosci. Remote Sens.* **51**(4), 2260–2275 (2013).
22. V. V. Lukin et al., "Denoising efficiency for multichannel images corrupted by signal-dependent noise," in *International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves*, Kharkiv, Ukraine, Vol. 3, pp. 340–342 (2013).
23. X. Liu, S. Bourennane, and C. Fossati, "Nonwhite noise reduction in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.* **9**(3), 368–372 (2012).
24. Y. Qian and M. Ye, "Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **6**(2), 499–515 (2013).
25. A. N. Zemliachenko et al., "VST-based lossy compression of hyperspectral data for new generation sensors," *Proc. SPIE* **8892**, 88920L (2013).
26. V. Lukin et al., "Analysis of classification accuracy for pre-filtered multichannel remote sensing data," *J. Expert Syst. Appl.* **40**(16), 6400–6411 (2013).
27. N. Acito, M. Diani, and G. Corsini, "Signal-dependent noise modeling and model parameter estimation in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.* **49**(8), 2957–2971 (2011).
28. M. Lebrun et al., "Secrets of image denoising cuisine," *Acta Numer.* **21**(1), 475–576 (2012).
29. S. Abramov et al., "Blind estimation of signal-dependent noise parameters for color image database," in *4th European Workshop on Visual Information Processing*, Paris, France, pp. 82–87 (2013).
30. C. Liu et al., "Noise estimation from a single image," in *Proc. IEEE Conf. on Computer Vision and Patter Recognition*, New York, Vol. 1, pp. 901–908 (2006).
31. M. L. Uss et al., "Image informative maps for component-wise estimating parameters of signal-dependent noise," *J. Electron. Imaging* **22**(1), 013019 (2013).
32. A. Foi, "Pointwise Shape-Adaptive DCT Image Filtering and Signal-Dependent Noise Estimation," Ph.D. Thesis, Tampere University of Technology, Tampere, Finland (2007).
33. F. J. Anscombe, "The transformation of poisson, binomial and negative-binomial data," *Biometrika* **35**, 246–254 (1948).
34. V. Lukin et al., "Lossy compression of images corrupted by mixed Poisson and additive Gaussian noise," in *International Workshop on Local and Non-Local Approximation in Image Processing*, pp. 33–40, Springer, Tuusula, Finland (2009).
35. F. Murtagh, "Astronomical image and signal processing," *IEEE Signal Process. Mag.* **18**(2), 30–40 (2001).
36. V. V. Lukin et al., "An automatic approach to lossy compression of images corrupted by Poisson noise," in *Microwaves, Radar and Remote Sensing Symposium*, Kiev, Ukraine, pp. 139–142 (2008).
37. A. Bovik, *Handbook of Image and Video Processing (Communications, Networking and Multimedia)*, pp. 1429, Academic Press, Inc., Orlando, Florida (2005).

38. E. Christophe, C. Mailhes, and P. Duhamel, "Hyperspectral image compression: adapting SPIHT and EZW to anisotropic 3-D wavelet coding," *IEEE Trans. Image Process.* **17**(12), 2334–2346 (2008).

39. L. Santos et al., "Performance evaluation of the H.264/AVC video coding standard for lossy hyperspectral image compression," *IEEE J. Sel. Topics Appl. Earth Observ.* **5**(2), 451–461 (2012).

40. N. Ponomarenko et al., "Improved grouping and noise cancellation for automatic lossy compression of AVIRIS images," in *Proc. Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, Sydney, Australia, Lecture Notes in Computer Science, J. Blanc-Talon et al., Eds., Vol. 6475, pp. 261–271, Springer Berlin Heidelberg (2010).

41. N. Ponomarenko et al., "Two aspects in lossy compression of hyperspectral AVIRIS images," in *12th International Conference on Mathematical Methods in Electromagnetic Theory*, Kharkiv, Ukraine, pp. 375–377 (2008).

42. V. Lukin, "Processing of multichannel RS data for environment monitoring," in *GeoSpatial Visual Analytics: Geographical Information Processing and Visual Analytics for Environmental Security*, NATO Science for Peace and Security Series C: Environmental Security, pp. 129–138, Springer Science and Business Media B.V., Berlin, Germany (2009).

43. M. L. Uss et al., "Image informative maps for component-wise estimating parameters of signal-dependent noise," *J. Electron. Imaging* **22**(1), 013019 (2013).

44. http://eros.usgs.gov/://eros.usgs.gov/.

45. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2004*, Pacific Grove, California, Vol. 2, pp. 1398 (2003).

46. N. Ponomarenko et al., "On between-coefficient contrast masking of DCT basis functions," in *Proc. Third Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scottsdale, Arizona, p. 4 (2007).

47. N. N. Ponomarenko et al., "DCT based high quality image compression," in *Proc. 14th Scandinavian Conf. on Image Analysis*, Joensuu, Finland, H. Kalviainen et al., Eds., pp. 1177–1185, Springer-Verlag, Berlin, Heidelberg (2005).

48. N. Ponomarenko et al., "Lossy compression of images with additive noise," in *Proc. Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, Antwerpen, Belgium, J. Blanc-Talon et al., Eds., pp. 381–386, Springer-Verlag, Berlin, Heidelberg (2005).

49. N. Ponomarenko et al., "Estimation of accessible quality in noisy image compression," in *14th European Signal Processing Conference*, pp. 4, Eurasip, Florence, Italy (2006).

50. V. Lukin et al., "Quasi-optimal compression of noisy optical and radar images," *Proc. SPIE* **6365**, 63650N (2006).

51. C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, SciTech Publishing, Raleigh, NC, pp. 486 (2004).

52. N. Ponomarenko et al., "ADCT: a new high quality DCT based coder for lossy image compression," in *Proc. Int. Workshop on Local and Non-Local Approximation in Image Processing*, pp. 1–6, Springer, Lausanne, Switzerland (2008).

53. A. Fidler, U. Skaleric, and B. Likar, "The impact of image information on compressability and degradation in medical image compression," *Med. Phys.* **33**(8), 2832–2838 (2006).

54. V. Lukin et al., "Lossy compression of images without visible distortions and its applications," in *10th International Conference on Signal Processing*, Beijing, China, pp. 698, 701 (2010).

55. N. Ponomarenko et al., "Performance analysis of visually lossless image compression," in *Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, CD-ROM*, Scottsdale, Arizona, p. 6 (2012).

56. J. H. Gruninger, A. J. Ratkowski, and M. L. Hoke, "The sequential maximum angle convex cone (SMACC) endmember model," *Proc. SPIE* **5425**, 1–14 (2004).

57. J.M.P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **43**(4), 898–910 (2005).

58. J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **5**(2), 354–379 (2012).

Alexander N. Zemliachenko received his BS degree in telecommunication engineering and his MS degree in electronics and telecommunications at National Aerospace University, Kharkov, Ukraine, in 2012. Currently, he is approaching his PhD in the area of image processing for remote sensing at the Department of Transmitters, Receivers and Signal Reception, National Aerospace University, Ukraine. He is the author of 7 research papers and 14 papers in the proceedings of international scientific conferences.

Ruslan A. Kozhemiakin received his BS and MS degrees in electronics and telecommunications engineering from National Aerospace University Kharkov, Ukraine, in 2012. He is a PhD candidate in the Department of Transmitters, Receivers, and Signal Processing at National Aerospace University. His current research interests include digital image compression and filtering in SAR and optical imagery.

Mikhail L. Uss graduated from National Aerospace University, Ukraine, in 2002 and has received a Diploma with honor in radio engineering. He obtained a Candidate of Technical Science degree in DSP for remote sensing from National Aerospace University, Ukraine, in 2006 and then his PhD degree from the University of Rennes 1, France, in 2011. His research interests include statistical theory of radio-technical systems, digital signal/image processing, fractal sets with applications to image processing, and remote sensing.

Sergey K. Abramov graduated from National Aerospace University in 2000 with a Diploma with honor in radio engineering. He defended his thesis of Candidate of Technical Science in 2003 in DSP for remote sensing. Currently, he is an associate professor and part-time senior researcher with the Department of Transmitters, Receivers, and Signal Processing. His research interests include digital signal/image processing, blind estimation of noise characteristics, and image filtering.

Nikolay N. Ponomarenko is an associate professor and senior researcher of the Department of Transmitters, Receivers, and Signal Processing of National Aerospace University of Ukraine. He obtained the Candidate of Technical Sciences degree from Ukraine in 2004, a Doctor of Technology degree from Tampere University of Technology, Finland, in 2005, and a Doctor of Technical Sciences degree (Ukraine) in 2012. He is the author of more than 200 scientific papers. His research interests include digital image processing, data compression, and data clustering.

Vladimir V. Lukin graduated from National Aerospace University in 1983, with his Diploma with honor in radio engineering. He defended the thesis of Candidate of Technical Science in 1988 and Doctor of Technical Science in 2002 in DSP for remote sensing. Since 1995, he has been in cooperation with Tampere University of Technology. Currently, he is department vice chairman and professor. His research interests include digital signal/image processing, remote sensing data processing, image filtering, and compression.

Benoît Vozel received his State Engineering degree and his MSc degree in control and computer science from École Centrale de Nantes, France, in 1991, and then his PhD degree from École Centrale de Nantes, France, in 1994. Since 1995, he has been with the Engineering School of Applied Sciences and Technology (ENSSAT), University of Rennes 1, France. His research interests generally concern blind estimation of noise characteristics, image restoration, adaptive image, and remote sensing data processing.

Kacem Chehdi (M'95) received his PhD and "Habilitation à diriger des recherches" degrees in signal processing and telecommunications from the University of Rennes 1, Rennes, France, in 1986 and 1992, respectively. He has been a professor of signal and image processing with the Engineering School of Applied Sciences and Technology (ENSSAT) since 1993. His research interests include adaptive processing at every level in the pattern recognition chain by vision, hyperspectral image processing, and analysis.