Global rate-distortion optimization of satellite imaging chains

Mikael Carlavan, Laure Blanc-Féraud, Marc Antonini, Carole Thiebaut, Christophe Latry, Yves Bobichon

GLOBAL RATE-DISTORTION OPTIMIZATION OF SATELLITE IMAGING CHAINS

M. Carlavan[1]*, L. Blanc-Féraud*[1]*, M. Antonini*[1]*, C. Thiebaut*[2]*, C. Latry*[2]*, Y. Bobichon*[3]

[1]I3S, 2000 route des Lucioles, 06902 Sophia-Antipolis, France
{carlavan, blancf, am}@i3s.unice.fr

[2]CNES, 18 avenue Edouard Belin, 31401 Toulouse, France
{Carole.Thiebaut, Christophe.Latry}@cnes.fr

[3]Thales Alenia Space, 100 boulevard du Midi, 06156 Cannes la Bocca, France
Yves.Bobichon@thalesaleniaspace.com

1. INTRODUCTION

The joint compression/restoration optimization of a satellite imaging chain is a challenging problem which has been little investigated so far. Some works have been done in designing an optimal coding/decoding structure which takes into account the characteristics of the imaging chain [1], but, to the best of our knowledge, the study of the global system optimization has not devoted much work; so that each process is usually optimized separately.

Images acquired by imaging systems are most of the time degraded by blur and noise, which mainly comes from the imperfections of the optical instrument and the electronic chain. It is well-known that noise decreases the performances of coding schemes as it reduces the correlation between pixels [2]. This problem is commonly referred as the noisy source coding problem [3]. Many works have been devoted to address this issue [4, 5, 6] and [7]. The majority of the mentioned works are based on the study of the global distortion optimization initially proposed by [5]. This study states that the global distortion, if measured by the Mean Square Error (MSE), can be treated as two separated problems. First, the original source image should be optimally, in the minimum MSE sense, estimated from the degraded data and this estimate should then be optimally coded [5].

However, an optimal estimation of the original image from the acquired one is usually not available. A suboptimal solution consists then in restoring the acquired image, before coding, and to encode this restored image [7]. Although we will also consider the imaging chain where the restoration is performed on-board before coding, critical applications such as satellite imaging may not afford to insert this supplementary step due to the high limitation of on-board resources. Therefore, one has to focus on the current imaging chain which restores the image after coding, and the global distortion of this chain needs then to be modeled and optimized. This is the focus of this paper.

We are considering here the problem of optimal coding and restoration of a degraded image. We show that, under certain hypotheses that we will describe, a closed-form expression of the global distortion can be obtained. We propose then to optimize this distortion, with respect to the coding and the restoration parameters, to reach the minimal global distortion. The originality of the proposed approach relies on the fact that we propose a global optimization which takes into account all the parameters of the imaging chain. More precisely, we emphasize the necessity to take into account the restoration step in the rate-distortion allocation of the coder. And, as we will see, this requirement is confirmed by the results which display a significant improvement in comparison to the classical method which executes coding and restoration independently. The analysis of the global distortion allows us to also address the question of the position of the restoration in the chain (on-board versus on-ground) by comparing numerically the minimal distortion obtained in each case.

The paper is organized as follows. In section 2, we present the studied imaging chain and introduce hypotheses and notations. We detail in section 3 the proposed approach and we show how to get a closed-form expression of the global distortion for the studied case. We detail the optimization of this distortion in section 4 and we present the algorithm to get the optimal parameters of the chain. We show results, visually and in a rate-distortion sense, of the proposed optimization algorithm on a remote sensing image, in section 5. We conclude in section 6 and present perspectives for future works.

2. HYPOTHESES AND NOTATIONS

In the following, the operators applied to the image are denoted with a bold uppercase letter. The non-bold uppercase letters represent random variables whose realizations are denoted by a lowercase letter. With this notation, x is a realization of the random variable X. $(X)_i$ denotes the i-nth element of the random variable X. These variables are multidimensional $x \in \mathbb{R}^N$ where N is the number of pixels. W_x is a random variable associated to the wavelet transform of x and we note $W_{x,j}, j \in \{0, \ldots, J-1\}$ (J being the number of subbands) the j-nth subband of the random variable W_x. We have $w_{x,j} \in \mathbb{R}^{N_j}$ where N_j is the size of the subband. Finally, we suppose that a wavelet subband $w_{x,j}$ follows a generalized centered Gaussian distribution law of parameter $\alpha_{w_{x,j}}$ and variance $\sigma^2_{w_{x,j}}$ [8]. The parameters $\sigma^2_{w_{x,j}}$ and $\alpha_{w_{x,j}}$ of the distribution law will be estimated using the kurtosis-based technique proposed in [9]. Note that the same assumption will be applied to all wavelet transforms in the chain.

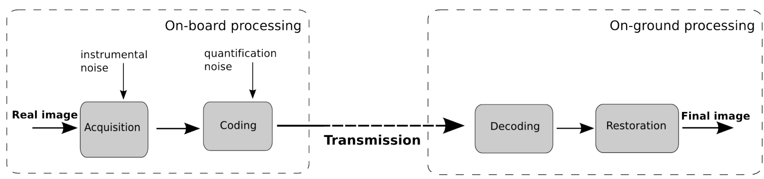

Fig. 1. On-ground satellite imaging chain.

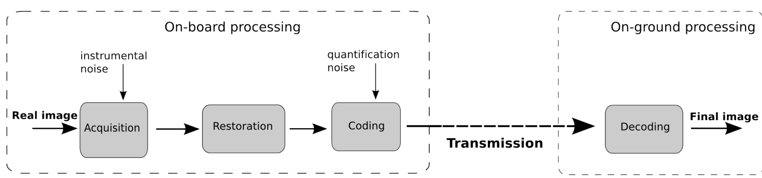

Fig. 2. On-board satellite imaging chain.

In this paper, we study the imaging chains shown figure 3 and 4. The imaging chain represented figure 3 is the usual satellite imaging chain which performs the restoration on-ground after the transmission of the image. In contrast with this chain, we also study the imaging chain depicted figure 4 which restores the image on-board before the coding step.

We consider the special case of coding techniques based on wavelet transforms [8]. The coding step is then approximately decomposed in a non-redundant wavelet transform followed by a scalar subband quantizer. We also consider that the denoising step is performed in the same wavelet basis than the coding. This choice may however need further explanations. Usually, an efficient wavelet transform for image denoising strongly differs from a wavelet transform suited for image coding. Image denoising techniques actually require redundant wavelet transforms to represent the characteristics of an image such as contours and oriented details while increasing the number of coefficients in image compression may be problematic [10]. Hence, a non-redundant wavelet transform used for image compression leads most of the time to poor denoising results. We are however very confident that using the same basis for both coding and denoising may provide an optimized coding-denoising structure gathered in a single fast and low-ressources algorithm. Extending the current work to complex denoising schemes such as [11] is a difficult task that will be addressed in future works.

Based on these considerations, the studied imaging chains are represented in detail on figures 3 and 4. Due to the space limitation, we only detail the proposed method for the on-ground chain shown figure 3. However, the extension of this work to the on-board chain displayed figure 4 is trivial and we will directly present the obtained results.
In the chain displayed figure 3, we consider the instrumental noise z to be independent, identically distributed and to follow a centered normal distribution with variance σ_z^2. $\mathbf{W}$ is the Cohen-Daubechies-Feauveau (CDF) 9/7 wavelet transform [12], $\tilde{\mathbf{W}}$ its inverse. Each quantized subband $w_{\tilde{y},j}$ will be coded using an entropy encoder. As this operation does not introduce any

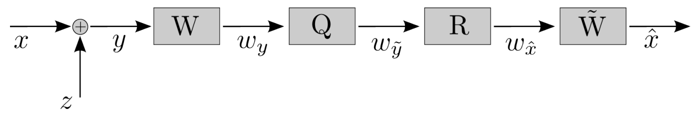

Fig. 3. On-ground considered imaging chain

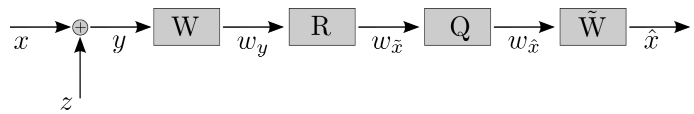

Fig. 4. On-board considered imaging chain

degradation, it does not appear on the chain. $\mathbf{R}$ is a linear restoration algorithm which operates on the wavelet coefficients of the image and writes

$$w_{\hat{x},j} = \begin{array}{rl} \arg\min & \|w - w_{\tilde{y},j}\|_2^2 + \lambda_j \|w\|_2^2 \\ \text{subject to} & w \in \mathbb{R}^{N_j} \end{array}, \tag{1}$$

where $\lambda_j > 0$ is a regularizing parameter. The restoration algorithm (1) has a closed-form solution which writes

$$w_{\hat{x},j} = \frac{w_{\tilde{y},j}}{1+\lambda_j}. \tag{2}$$

We are aware of the simplicity of the considered algorithm, it appears however that the linearity of the restoration algorithm $\mathbf{R}$ is required if one wants to write the global distortion in closed-form. As mentioned previously, much work need to be addressed to consider state-of-art denoising algorithms.

The quantizer $\mathbf{Q}$ is an infinite scalar subband quantizer of step $\Delta_j > 0$ and is modeled as

$$\mathbf{Q}(w_{y,j}) = \Delta_j \left\lfloor \frac{w_{y,j}}{\Delta_j} + \frac{1}{2} \right\rfloor, \tag{3}$$

where $\lfloor\ \rfloor$ is the floor function which returns the greatest integer less than or equal to its argument. We now detail the basis of the proposed method to compute a closed-form expression of the global distortion.
Let $w_{b,j}$ be the coding error of the subband j

$$w_{b,j} = \mathbf{Q}(w_{y,j}) - w_{y,j}. \tag{4}$$

We have

$$w_{\tilde{y},j} = w_{y,j} + w_{b,j} = w_{x,j} + w_{\epsilon,j}, \tag{5}$$

where $w_{\epsilon,j} = w_{z,j} + w_{b,j}$. The main hypothesis of the proposed method is to consider the moments of the term $w_{\epsilon,j}$ to be independent to the ones of $w_{x,j}$. This hypothesis is mainly based on the fact that the quantizing part of the scheme figure 3 can be seen as a non-substractive dithering system where the Gaussian instrumental noise acts as a dithering noise. And, as mentioned in [13], a non-substractive dithering system allows the moments of the global error (that is the sum of the coding and instrumental errors) to be decorrelated to the moments of the coding source. More precisely, we have [13]

$$E\left[W_{\epsilon,j} W_{x,j}\right] = E\left[W_{\epsilon,j}\right] E\left[W_{x,j}\right], \tag{6}$$

$$E\left[W_{\epsilon,j}\right] = 0, \tag{7}$$

$$E\left[\|W_{\epsilon,j}\|^2\right] = N_j \sigma_{w_{z,j}}^2 + N_j \frac{\Delta_j^2}{12}, \tag{8}$$

where $\sigma_{w_{z,j}}$ is the standard deviation of the distribution law of the wavelet transform $w_{z,j}$. From [14] we know that a Gaussian noise effectively owns the properties of a dither noise if the standard deviation of its distribution law is large enough. In the present case, this dithering hypothesis will be verified if the following statement is true

$$\sigma_{w_{z,j}} > \frac{\Delta_j}{2}. \tag{9}$$

As the standard deviation of instrumental noise is usually low in imaging systems, the condition (9) assumes that the proposed approach will be valid only for aymptotic coding rates, i.e. high coding rates. We will however develop our method to consider all coding rates.

3. GLOBAL RATE-DISTORTION ANALYSIS

As mentioned in the section 2, the studied imaging chain depends on two sets of parameters: The regularizing parameters λ_j in (2) and the quantizing steps Δ_j in (3), for each $j \in \{0, \ldots, J-1\}$. The global rate-distortion optimization problem consists in finding the optimal parameters λ_j^* and Δ_j^* which minimize the global distortion D under the constraint that the coding rate R does not exceed the target rate R_c. This can be formalized as the following unconstrained problem

$$\text{Find} \quad \lambda_j^*, \Delta_j^* = \begin{array}{rl} \arg\min & D(\lambda_j, \Delta_j) + \tau\left(R(\lambda_j, \Delta_j) - R_c\right) \\ \text{subject to} & \lambda_j > 0, \Delta_j > 0, \tau > 0 \end{array}, \tag{10}$$

where τ is a Lagrange multiplier. To solve the global distortion optimization problem (10), we then need to express the global distortion D and the global coding rate R as a function of the regularizing parameters λ_j and the quantizing steps Δ_j. We start from the fact that the (mean) global distortion writes

$$D = \frac{1}{N} E\left(\|X - \hat{X}\|^2\right), \tag{11}$$

where $\hat{X}$ is the random variable associated to the output final image $\hat{x}$. Thanks to the orthogonality of the wavelet subbands, the global distortion can also be formulated as

$$D = \frac{1}{N} \sum_{j=0}^{J-1} \pi_j E\left(\|W_{x,j} - W_{\hat{x},j}\|^2\right), \tag{12}$$

where π_j are weighting coefficients as the considered wavelet transform is bi-orthogonal [15]. In the case of the studied imaging chain displayed figure 3, the final image can be expressed as a function of the source and the global error. Using equations (2) and (5), it writes

$$w_{\hat{x},j} = \frac{w_{x,j}}{1+\lambda_j} + \frac{w_{\epsilon,j}}{1+\lambda_j}. \tag{13}$$

From equations (12), (13) and using the moments decorrelation hypothesis (6), we deduce the global distortion

$$D = \frac{1}{N} \sum_{j=0}^{J-1} \frac{\pi_j \lambda_j^2}{(1+\lambda_j)^2} E\left(\|W_{x,j}\|^2\right) + \frac{\pi_j}{(1+\lambda_j)^2} E\left(\|W_{\epsilon,j}\|^2\right). \tag{14}$$

Finally, the global distortion (14) can be further developed using the results (8)

$$D = \sum_{j=0}^{J-1} \frac{\pi_j a_j \lambda_j^2}{(1+\lambda_j)^2} \sigma_{w_{x,j}}^2 + \frac{\pi_j a_j}{(1+\lambda_j)^2} \sigma_{w_{z,j}}^2 + \frac{\pi_j a_j}{(1+\lambda_j)^2} \frac{\Delta_j^2}{12}, \tag{15}$$

where $a_j = \frac{N_j}{N}$ is the weight of the subband j in the whole image. Note that the global distortion (15) requires the knowledge of the variance of each subband of the original image $\sigma_{w_{x,j}}^2$. This variance is generally unknown but can be approximatively computed during the rate-allocation of the coder from the observed subband variance $\sigma_{w_{y,j}}^2$, as the variance of the instrumental noise is assumed to be known. The second part of the problem (10) requires the expression of the global coding rate R. This rate can be expressed as the weighted sum of the rate in each subband R_j

$$R = \sum_{j=0}^{J-1} a_j R_j(\Delta_j). \tag{16}$$

As mentioned in the hypotheses section, we assume that each quantized subband is encoded using an entropy encoder. Then, in a general non-asymptotic case, the coding rate of a subband j can be estimated by its entropy [16]

$$R_j(\Delta_j) = -\sum_{m=-\infty}^{+\infty} P_{w_{y,j}}(m, \Delta_j) \log_2\left(P_{w_{y,j}}(m, \Delta_j)\right), \tag{17}$$

where $P_{w_{y,j}}(m, \Delta_j)$ is the probability of a coefficient of the quantized subband $w_{\tilde{y},j}$ to be equal to the symbol m. The same analysis can be applied to the on-board chain figure 4 to modelize the global distortion. It writes

$$D = \sum_{j=0}^{J-1} \frac{\pi_j a_j \lambda_j^2}{(1+\lambda_j)^2} \sigma_{w_{x,j}}^2 + \frac{\pi_j a_j}{(1+\lambda_j)^2} \sigma_{w_{z,j}}^2 + \frac{\pi_j a_j \Delta_j^2}{12}. \tag{18}$$

We detail in the next part how to solve the optimization problem of the on-ground global rate-distortion.

4. GLOBAL RATE-DISTORTION OPTIMIZATION

Using equations (15) and (16), the optimization problem (10) becomes

$$\text{Find} \quad \Delta_j^*, \lambda_j^* = \begin{array}{ll} \arg\min & \phi(\Delta_j, \lambda_j, \tau) \\ \text{subject to} & \lambda_j > 0, \Delta_j > 0, \tau > 0 \end{array}. \tag{19}$$

where

$$\phi(\Delta_j, \lambda_j, \tau) = \sum_{j=0}^{J-1} \frac{\pi_j a_j \lambda_j^2}{(1+\lambda_j)^2} \sigma_{w_{x,j}}^2 + \frac{\pi_j a_j}{(1+\lambda_j)^2} \sigma_{w_{z,j}}^2 + \frac{\pi_j a_j \Delta_j^2}{12(1+\lambda_j)^2} + \tau \left(\sum_{j=0}^{J-1} a_j R_j(\Delta_j) - R_c \right). \tag{20}$$

The theoretical study of existence and uniqueness of solutions of problem (19) is an open problem and we explicitly assume that a minimum of problem (19) exists and is unique. We propose a numerical algorithm to find this minimum. This algorithm is based on the resolution of the simultaneous equations obtained from the Karush-Kuhn-Tucker (KKT) conditions [17] of problem (19). The KKT conditions of problem (19) write

$$\begin{cases} \frac{\partial \phi(\Delta_j^*, \lambda_j^*, \tau^*)}{\partial \Delta_j} = \frac{\pi_j \Delta_j^*}{6(1+\lambda_j^*)^2} + \tau^* \frac{\partial R_j}{\partial \Delta_j}(\Delta_j^*) = 0 \\ \frac{\partial \phi(\Delta_j^*, \lambda_j^*, \tau^*)}{\partial \tau} = \sum_{j=0}^{J-1} a_j R_j(\Delta_j^*) - R_c = 0 \\ \frac{\partial \phi(\Delta_j^*, \lambda_j^*, \tau^*)}{\partial \lambda_j} = \frac{12\lambda_j^* \sigma_{w_{x,j}}^2 - 12\sigma_{w_{z,j}}^2 - \Delta_j^{*2}}{6(1+\lambda_j^*)^3} = 0 \end{cases} \tag{21}$$

with

$$\frac{\partial R_j}{\partial \Delta_j}(\Delta_j) = -\frac{1}{\log(2)} \sum_{m=-\infty}^{+\infty} \left[1 + \log\left(P_{w_{y,j}}(m, \Delta_j)\right)\right] \times$$
$$\left[p_{w_{y,j}}\left(m\Delta_j + \frac{\Delta_j}{2}\right)\left(m + \frac{1}{2}\right) - p_{w_{y,j}}\left(m\Delta_j - \frac{\Delta_j}{2}\right)\left(m - \frac{1}{2}\right)\right]. \tag{22}$$

We deduce from (21) that the optimization problem (19) admits a solution (λ_j^*, τ^*, Δ_j^*) which verifies [18]

$$\lambda_j^* = \frac{\sigma_{w_{z,j}}^2}{\sigma_{w_{x,j}}^2} + \frac{\Delta_j^{*2}}{12\sigma_{w_{x,j}}^2}, \tag{23}$$

$$\frac{\pi_j \Delta_j^*}{6(1+\lambda_j)^2} + \tau^* \frac{\partial R_j}{\partial \Delta_j}(\Delta_j^*) = 0, \tag{24}$$

$$\sum_{j=0}^{J-1} a_j R_j(\Delta_j^*) = R_c. \tag{25}$$

As shown by equations (24) and (25), the optimal quantizing step Δ_j^* and the optimal Lagrange multiplier τ^* can not be expressed in closed-form. These parameters can however be computed using any root-finding algorithm [18]. The case of the low frequency subband ($j = J-1$) will be treated differently as we do not quantize it. We will set $\Delta_{J-1}^* = 1$ and will deduce λ_{J-1}^* from (23). The overall optimization procedure for solving problem (10) is given in the algorithm 1. The algorithm 1 intends to be quite general and we let the choice of the root-finding algorithm to the user (we note ρ the precision of this algorithm).

5. RESULTS

We simulate the optimization algorithm 1 on the high-dynamic range remote sensing image displayed figure 5. For this simulation, the reference image has been noised with an additive centered Gaussian noise with a standard deviation equal to 10. For each target rate, we simulate the imaging chains figures 3 and 4 with the common choice of selecting the quantizing steps and the regularizing parameters such that the coding and the restoration errors are independently minimized. The coding error minimization has been achieved using the rate-distortion allocation based model proposed in [1]. As for the restoration error, it has

Algorithm 1 Global rate-distortion optimization algorithm

Set $\tau = 1$ and $\rho = 0.1$.
while $\left|\sum_{j=0}^{J-1} a_j R_j - R_c\right| > \rho$ **do**
  for j from 0 to $J-2$ **do**
    Set $\Delta_j = 1$.
    Compute the value of the regularizing parameter λ_j from (23).
    while $\left|\frac{\pi_j \Delta_j}{6(1+\lambda_j)^2} + \tau \frac{\partial R_j}{\partial \Delta_j}(\Delta_j)\right| > \rho$ **do**
      Increase the value of Δ_j.
      Compute the value of the regularizing parameter λ_j from (23).
    end while
  end for
  Set $\Delta_{J-1} = 1$.
  Compute the regularizing paramater λ_{J-1} from (23).
  if $\left|\sum_{j=0}^{J-1} a_j R_j - R_c\right| > \rho$ **then**
    Increase the value of τ.
  end if
end while
Output the optimal regularizing parameters λ_j^* and the optimal quantizing steps Δ_j^*.

been minimized using an exhaustive search of the optimal regularizing parameters. Once the final image has been reconstructed using these parameters, we compute numerically the global distortion

$$D = \frac{1}{N}\|x - \hat{x}\|^2, \tag{26}$$

where x is the reference image (assumed to be known in our numerical experiments) and $\hat{x}$ is the final image. The proposed model of distortion (15) has then been computed with the same values of parameters to ensure its validity. And finally, we use the proposed algorithm 1 to compute the joint optimization. The obtained parameters have then been inserted into the theoretical model of distortion (15) to compute the minimal distortion. The same comparison has been applied to the on-board imaging chain displayed figure 4.

The resulting rate-distortion curves are given figure 5. We see that the proposed model approximates well the true distortion if the coding rate is around 3 bits/pixel and more, see figure 5. It is however not valid below this rate. This can explained by the fact that below this rate, the condition (9) is not respected anymore and that the global error can not consequently be considered decorrelated to the source. This bound may be improved by considering a more realistic model for the variance of the global error (8) rather than the classical $\frac{\Delta_j^2}{12}$ term used for asymptotic rate hypotheses [19]. Even if it does not solve the correlation issue, refining the model may allow to fit the true distortion for lower rates. This will be the subject of future works. Regarding to the on-board/on-ground comparison, it seems from the results that the on-ground chain always gives better results, at least for the studied case. It should be noted however that, in the case of the on-board chain, the coding noise is not processed. An "hybrid" chain, with one restoration on-board and one on-ground, may then give interesting results. This will also be the suject of future works.

Visual results for the target rate of 2.5 bits/pixel are given figure 6 for different parts of the image. We do not focus on the quality of the reconstructed images regarding to the reference one as the considered chain is excessively simple. Clearly, the presence of artifacts on the reconstructed image is due to the simple hypothesis that we made on the restoration algorithm (1). On the contrary, we are more concerned on the improvement of the image quality of the optimized chain with respect to the non-optimized one. We can see that the global optimization of the chain always lead to a reconstructed image witch contains less blurry edges or ringing artifacts. This is particularly visible on the boundaries of objects (cars for example). As mentioned in the introduction of this paper, the obtained results clearly point that optimizing coding and restoration separately is suboptimal. One needs instead to address the problem of imaging chain design in its globality; the proposed method and the obtained results are encouraging in this sense. A lot of works is however required to extend the proposed method to lower coding rates and to more complex restoration techniques.

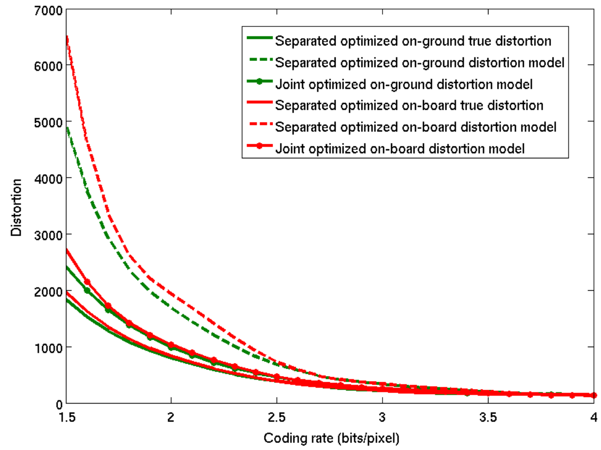

Fig. 5. On the left: Reference image (12 bits), Cannes harbour (1024×1024 pixels). On the right: Comparison of the separated optimized distortion model and its joint optimization to the true separated optimized distortion on the image of *Cannes*, for each (on-board and on-ground) imaging chain

6. CONCLUSIONS

In this paper we considered the problem of satellite imaging chain optimization. Most of the time, the coding and the restoration parameters are selected independently such that the coding and the restoration error are respectively minimized. As shown in this paper, this parameters selection technique leads however to a suboptimal distortion. It appears then crucial to address the problem of joint coding/denoising in its globality. We proposed here a technique to modelize the global distortion and we presented an algorithm to get the optimal coding and denoising parameters. We simulated this optimization technique on a high-dynamic range remote sensing image. We concluded that our joint coding/denoising optimization approach really improves the quality of the reconstructed final image, in comparison to the image obtained using the classical parameters selection technique. Further works will be focussed on the extension of the proposed model to lower coding rates and to advanced restoration methods.

7. REFERENCES

[1] C. Parisot, M. Antonini, M. Barlaud, S. Tramini, C. Latry, and C. Lambert-Nebout, "Optimization of the joint coding/decoding structure," in *International Conference on Image Processing (ICIP)*, 2001.

[2] S.-C.B. Lo, B. Krasner, and S.K. Mun, "Noise impact on error-free image compression," *IEEE Transactions on Medical Imaging*, vol. 9, no. 2, pp. 202–206, Jun. 1990.

[3] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 293–304, Sep. 1962.

[4] Y. Ephraim and R.M. Gray, "A unified approach for encoding clean and noisy sources by means of waveform and autoregressive model vector quantization," *IEEE Transactions on Information Theory*, vol. 34, no. 4, pp. 826–834, Jul. 1988.

[5] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Transactions on Information Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.

[6] T.R. Fischer, J.D. Gibson, and B. Koo, "Estimation and noisy source coding," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 1, pp. 23–34, jan 1990.

[7] O.K. Al-Shaykh and R.M. Mersereau, "Lossy compression of noisy images," *IEEE Transactions on Image Processing*, vol. 7, no. 12, pp. 1641–1652, Dec. 1998.

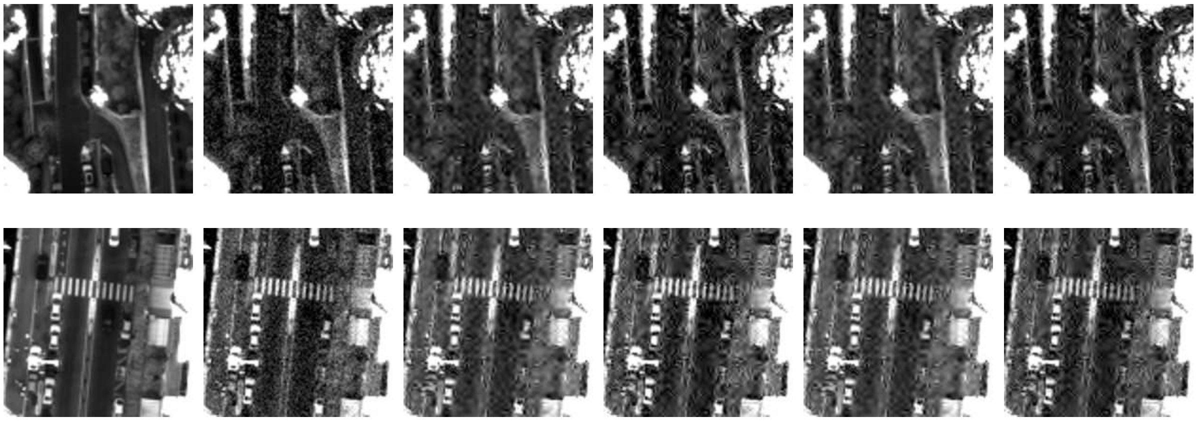

Fig. 6. Visual comparison of reconstruction results. From left to right: Reference image, observed image, image reconstructed using the separated optimization of the on-ground distortion, image reconstructed using the joint optimization of the on-ground distortion model, image reconstructed using the separated optimization of the on-board distortion and image reconstructed using the joint optimization of the on-board distortion model. The coding rate is 2.5 bits/pixel. The image range has been extended to point up the image reconstruction artifacts.

[8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, Apr. 1992.

[9] J.H. Kasner, M.W. Marcellin, and B.R. Hunt, "Universal trellis coded quantization," *IEEE Transactions on Image Processing*, vol. 8, no. 12, pp. 1677–1687, 1999.

[10] V. Chappelier and C. Guillemot, "Oriented wavelet transform for image compression and denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 2892–2903, 2006.

[11] D.L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[12] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 45, no. 5, pp. 485–560, 1992.

[13] R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, and J.N. Wright, "A theory of nonsubtractive dither," *IEEE Transactions on Signal Processing*, vol. 48, no. 2, pp. 499–516, Feb. 2000.

[14] J. Vanderkooy and S.P. Lipshitz, "Dither in digital audio," *Journal of the Audio Engineering Society*, vol. 35, no. 12, pp. 966–975, Feb. 1987.

[15] B. Usevitch, "Optimal bit allocation for biorthogonal wavelet coding," in *Data Compression Conference*, Mar. 1996, pp. 387–395.

[16] C.E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, 1948.

[17] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Berkeley Symposium on Mathematical Statistics and Probability*, 1951, pp. 481–492.

[18] M. Carlavan, L. Blanc-Féraud, M. Antonini, C. Thiebaut, C. Latry, and Y. Bobichon, "An optimized structure for joint noisy image coding/denoising," Tech. Rep., 2012.

[19] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, no. 5, pp. 676–683, Sep. 1968.