applied sciences

MDPI

Article

BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters

Bogdan Kovalenko [1], Vladimir Lukin [1,*], Sergii Kryvenko [1], Victoriya Naumenko [1] and Benoit Vozel [2]

[1] Department of Information and Communication Technologies, National Aerospace University, 61070 Kharkiv, Ukraine; b.kovalenko@khai.edu (B.K.); krivenkos@ieee.org (S.K.); v.naumenko@khai.edu (V.N.)
[2] IETR, UMR CNRS 6164, University of Rennes, 22305 Lannion, France; benoit.vozel@univ-rennes1.fr
* Correspondence: v.lukin@khai.edu

Abstract: With a resolution improvement, the size of modern remote sensing images increases. This makes it desirable to compress them, mostly by using lossy compression techniques. Often the images to be compressed (or some component images of multichannel remote sensing data) are noisy. The lossy compression of such images has several peculiarities dealing with specific noise filtering effects and evaluation of the compression technique's performance. In particular, an optimal operation point (OOP) may exist where quality of a compressed image is closer to the corresponding noise-free (true) image than the uncompressed (original, noisy) image quality, according to certain criterion (metrics). In such a case, it is reasonable to automatically compress an image under interest in the OOP neighborhood, but without having the true image at disposal in practice, it is impossible to accurately determine if the OOP does exist. Here we show that, by a simple and fast preliminary analysis and pre-training, it is possible to predict the OOPs existence and the metric values in it with appropriate accuracy. The study is carried out for a better portable graphics (BPG) coder for additive white Gaussian noise, focusing mainly on one-component (grayscale) images. The results allow for concluding that prediction is possible for an improvement (reduction) in the quality metrics of PSNR and PSNR-HVS-M. In turn, this allows for decision-making about the existence or absence of an OOP. If an OOP is absent, a more "careful" compression is recommended. Having such rules, it then becomes possible to carry out the compression automatically. Additionally, possible modifications for the cases of signal-dependent noise and the joint compression of three-component images are considered and the possible existence of an OOP for these cases is demonstrated.

Keywords: image lossy compression; optimal operation point; quality prediction; noise; discrete cosine transform; automation

check for updates

Citation: Kovalenko, B.; Lukin, V.; Kryvenko, S.; Naumenko, V.; Vozel, B. BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters. *Appl. Sci.* **2022**, *12*, 7555. https://doi.org/10.3390/app12157555

Academic Editors: Zhenglei He, Yi Man and Kim Phuc Tran

Received: 29 May 2022
Accepted: 26 July 2022
Published: 27 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Remote sensing (RS) systems and other imaging tools currently provide valuable data for agriculture, forestry, hydrology, ecological monitoring, non-destructive testing in industry, etc. [1–4]. Using imaging data produced by modern systems, it is possible to estimate the parameters of sensed territories of large areas, to control their change in time, to detect objects of interest and to solve many other important tasks. Meanwhile, a better spatial resolution and more frequent observation result in a fast increase in the data volume where the images have to be transferred, processed, stored and interpreted [5,6].

At the stages of data transferring and storage, image compression is often applied [7–9]. A lossless compression [7,8,10] preserves all the information contained in RS and imaging data, but the compression ratio (CR) attained for the used methods can be inappropriate in practice. To obtain a larger CR, near-lossless and lossy compression is mostly applied [7,9–12]; however, distortions are then inevitably introduced and the important task

of providing an appropriate trade-off between the introduced distortions and the reached CR arises [13,14]. The priority of requirements depends upon the application at hand and it can be necessary or desired to provide a given CR (or a CR not less than a given threshold). Meanwhile, it can also be necessary (or desired) to provide a desired quality (or a quality not worse than desired), but there can also be other requirements or restrictions that might concern the necessity for relying on standards or to reach a trade-off quickly, to save resources and to carry out a compression fully automatically, etc. In this paper, we focus on providing high quality with the desire to have quite a high CR and to perform the compression quite quickly and automatically. If the quality degradation due to a lossy compression is limited, it is possible to expect that the RS data classification or object (e.g., crack) detection are performed well enough, and that the RS data being visualized are of a proper quality for analysis and so on [15–17].

In many cases, it is supposed that the images to be compressed are noise-free or, at least, that the noise is invisible. This is often valid in practice but, meanwhile, there are quite a number of practical situations when noise is visible and its presence cannot be neglected; for example, noise (speckle) is always seen in radar images [18,19], some component images of hyperspectral and multispectral RS data are noisy [20], the signal-to-noise ratio can be low in night-light images [21] or in images acquired in bad illumination conditions and so on. The influence of noise first attracted the attention of researchers more than 20 years ago [22,23]. It has been demonstrated that the lossy compression of noisy images has two main peculiarities. First, a specific noise filtering effect is observed. Second, due to this, a so-called optimal operation point (OOP) can be observed, i.e., such a parameter of a coder that the "distance" between the compressed and true (noise-free) images is minimal. By distance, we mean some similarity measure, and this can be a mean square error (MSE, then the OOP would correspond to its minimum) or the peak signal-to-noise ratio (PSNR, then the OOP would correspond to its maximum) [22–24]. The OOP can also be observed for visual quality metrics [24,25] such as PSNR-HVS-M [26] and MS-SSIM [27]. Note that the use of visual quality metrics has become popular in many modern applications including in stereoscopic, panoramic and 360 degree imaging [28–30].

Compression in the OOP (or its neighborhood if the OOP is determined with some error), if it exists, has two advantages. First, the provided CR is usually quite high. Second, the quality of the compressed image appears to be better than the quality of the uncompressed (original, noisy, or compressed in a lossless manner) image. This can be, e.g., favorable for image classification [31]. Additionally, note that the OOP might exist not only for additive noise but also for other types of noise including signal-dependent and multiplicative for methods that employ a variance stabilizing transform (VST) or that perform compression without VST. An OOP might also exist for different coders, based both on a discrete cosine transform (DCT) or wavelets (for example, JPEG2000) [32]. Recently, the possible existence of an OOP [33,34] has been demonstrated for the coder, BPG (better portable graphics} [35–37]. This compression technique has several advantages. It outperforms JPEG considerably in the sense of a better quality for a given size of compressed images and the compression is fast enough with clear rules of providing a desired PSNR. Due to this, the encoder has become popular in different portable devices as well as for online encoding applications and these advantages were the reasons why BPG has attracted our attention [36,37]. The current paper incorporates the results of [36] and concentrates on the prediction of the OOP's existence and estimating the performance parameters for it. This is explained by the following. Since the true image is absent, the OOP cannot be determined exactly and its existence or absence can only be predicted (at least, this has been shown for coders other than BPG coders) [24,29]. Assuming that it is also possible for BPG, it can then be supposed that it will be possible to give recommendations for how to set the compression parameters (namely, the quality parameter Q for BPG) to provide an appropriate quality. This is the main goal of this paper.

The paper contributions consist in the following. First, based on the results of many test images corrupted by AWGN having different values of noise variance, a Q for a possible

OOP is established depending on the noise standard deviation; this dependence is shown to be logarithmic opposite to the linear dependences established in [24] for other coders. Second, a modification of the procedure [24] is proposed and its thorough analysis is carried out. In particular, it is proposed to use rational functions in curve fitting into scatter-plots. This analysis shows that there are, at least, two statistical parameters that can be quickly and easily calculated in the DCT domain that can be used for predicting an improvement or reduction in two metrics. Third, the accuracy of prediction is analyzed and factors such as noise realization and the number of blocks that influence the accuracy are considered. Based on this analysis, practical recommendations are given that allows compression automation. Fourth, initial data concerning the lossy compression of single-channel images corrupted by Poisson noise and three-channel images corrupted by AWGN are presented, demonstrating the possible existence of OOPs for these cases. Finally, we show that the BPG coder is more efficient compared to other coders earlier considered in [24].

The paper is structured as follows. Section 2 describes the image/noise model, considered metrics and basic dependences. Section 3 analyzes the dependences more in detail. The methodology of prediction and its accuracy analysis are given in Section 4. Decision-making and practical recommendations are discussed in Section 5. The cases of signal-dependent noise and three-channel image compression are also briefly studied in this Section. Finally, the conclusions are given.

2. Image/Noise Model and Compression Efficiency Criteria

It is a well-known fact that compression characteristics depend on sufficient image properties. Depending on the image complexity, the CR can vary by several (or even tens) of times for the same PSNR or, equivalently, the quality of compressed images can be rather different for the same CR. Because of this, to ensure universality of the conclusions and recommendations, the corresponding analysis and method synthesis should be performed for a set of images with a very wide variation of properties. A set of test images needs to contain simple, medium, and complex structure images where the complexity can be described or characterized by a percentage of pixels that belong to an image's homogeneous regions or entropy for a noise-free case and so on. For a better understanding, we give six images as the examples of images of different complexity (Figure 1). More than a half of the pixels in Figure 1a belong to quasi-homogeneous regions, a certain part of the pixels in Figure 1c–f relate to homogeneous regions, whilst there are practically no homogeneous regions in the image in Figure 1b.

There are different ways to characterize image complexity. We have noticed that the performance of the lossy compression of noisy images correlates with the image entropy for noise-free cases. Entropy E is the following: 5.82 for Frisco, 7.33 for Diego, 7.46 for Fr01, 7.40 for Fr02, 7.38 for Fr03, and 7.29 for Fr04. It will be shown later that the plots for the test images having similar E values for the noise-free versions usually have a similar behaviour. Note that the marginal cases are of the most interest; however, "average" cases (of middle complexity) are also worth considering. Because of this, there is a tendency in the image processing community for the creation of image databases and for their use in the design and verification of image processing methods [38,39]. As a starting point for analyzing the lossy compression of noisy images using BPG, we concentrate on the case of grayscale images.

As a noise model, we consider the AWGN, which is known to be the simplest model. Just such a model is commonly used as a starting point in research, for example, in the prediction of the potential efficiency of image denoising [40]; thus, one can present an observed noisy grayscale image as:

$$I_{ij}^{n} = I_{ij}^{true} + n_{ij}, \quad (1)$$

where I_{ij}^{true}, $i = 1, ..., I_{Im}$, $j = 1, ..., J_{Im}$ is the true or noise-free image, n_{ij} denotes AWGN in the ij-th pixel, and I_{Im} and J_{Im} define the considered image size. It is possible to assume the

noise mean equals to zero and the AWGN variance is equal to σ^2. Moreover, we assume that σ^2 is either a priori known or pre-estimated with a high accuracy [41].

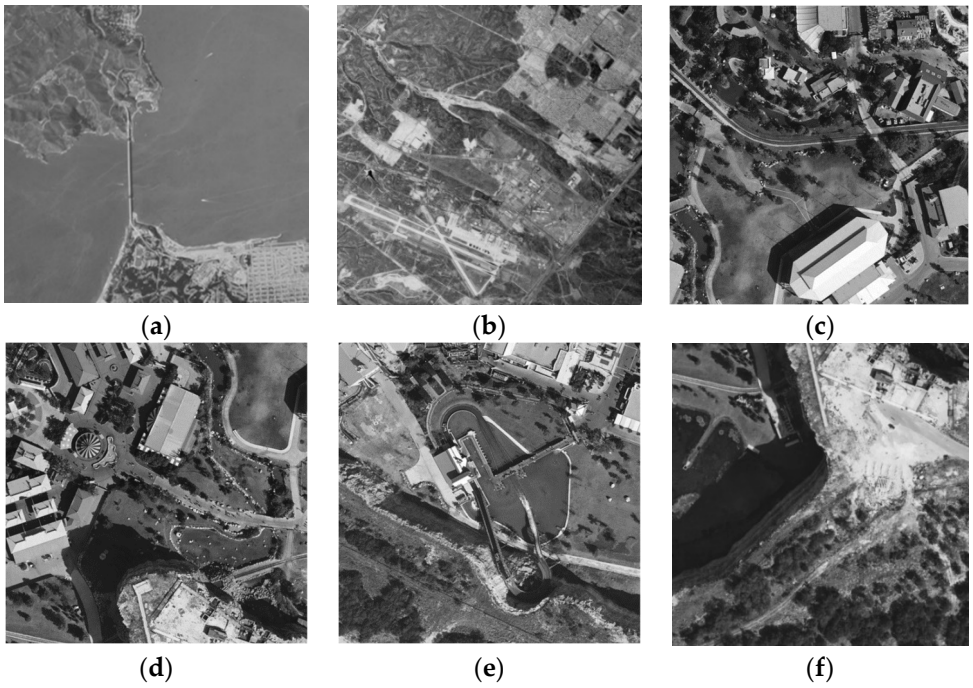


Figure 1. 512 × 512 pixel images having simple (**a**), complex (**b**), and middle complexity (**c**–**f**) structures (namely, Frisco (**a**), Diego (**b**), Fr01 (**c**), Fr02 (**d**), Fr03 (**e**), and FR04 (**f**)).

Quality of the original noisy image can be characterized in different ways. One standard way is to calculate the peak signal-to-noise ratio as:

$$PSNR^n = 10log_{10}\left(\frac{255^2}{\sigma^2}\right), \tag{2}$$

under the assumption that an image is represented as an 8-bit two-dimensional (2D) data array. Another way is to employ some visual quality metrics, here, we use two of them. The first one is PSNR-HVS-M [26] (a peak signal-to-noise ratio taking into account the human vision system (HVS) and masking (M)), and the multi-scale structural similarity metric (MS-SSIM) [27]. Both are among the best in characterizing the visual quality of images with distortions typical for remote sensing [42], in particular those due to noise and lossy compression. They can be calculated quickly, are applicable to a single-channel (grayscale images), and are based on different principles. The latter is important since no elementary full-reference metric is perfect and, thus, while carrying out an analysis and making conclusions, it is desired to rely on the results obtained for several visual quality metrics.

PSNR-HVS-M is defined as:

$$PSNR - HVS - M^n = 10log_{10}\left(\frac{255^2}{MSE - HVS - M^n}\right), \tag{3}$$

where $PSNR - HVS - M^n$ is determined in 8 × 8 blocks in the DCT domain, taking into account the masking effect and lower sensitivity of a human eye to distortions in high spatial frequencies rather than distortions in low spatial frequencies. Note that PSNR and PSNR-HVS-M [26] are both expressed in dB with larger values relating to better quality. Usually, for AWGN and similar distortions, PSNR-HVS-M is slightly larger than PSNR due to masking effect.

For these two metrics, it is important that the distortion visibility thresholds have been determined [38]. Distortions are usually invisible if PSNR exceeds 36 dB and PSNR-HVS-M is larger than 41 dB. This happens if the noise variance is larger than 15...20. Because of this, in this paper we concentrate on considering $PSNR^n$ smaller than 36 dB which is typical for the aforementioned applications.

The metric MS-SSIM [27] extracts and employs structural information from the scene. It is in the limits from zero to unity where larger values correspond to a better visual quality. The distortion invisibility threshold is approximately equal to 0.99.

If a lossy compression method is applied, one obtains a compressed image I^c_{ij}, $i = 1, ..., I_{Im}$, $j = 1, ..., J_{Im}$ that differs from I^n_{ij}, $i = 1, ..., I_{Im}$, $j = 1, ..., J_{Im}$ and depends on the compression controlling parameter (CCP). For different coders, different parameters play the role of the CCP. This can be the quality factor, scaling factor, quantization step, and bits per pixel [12,24,31,32]. For the BPG coder, the quality parameter Q allows changing the image quality and compression ratio. Performing a multiple compression of I^n using different Q (which are integers in the limits from 1 to 51 for the BPG encoder), the rate-distortion curve $Metr^{nc}(Q)$ can be obtained for any image where $Metr^{nc}$ is a metric under interest calculated between noisy (original) and compressed images.

Such metrics behave in a reasonable manner, i.e., their values become worse (smaller for all three considered metrics) if the Q increases. This is clearly seen in all four plots presented in Figure 2. In Figure 2a,b, the dependences of $PSNR^{nc}(Q)$ for six test images for two values of the noise variance are presented. For very small $Q < 7$, the original and compressed images practically do not differ and the compression is near-lossless. Then, one has a practically linear part of all curves; thus, it is possible to say that the part starts at $Q = 7$ and ends at such Q that $PSNR^{nc}(Q) \approx PSNR^n$. For larger Q, the dependences for different images diverge where larger $PSNR^{nc}$ values are observed for simpler structure images. For the middle part, it is possible to approximate the curves as:

$$PSNR^{nc}(Q) \approx 63 - Q, \; dB. \tag{4}$$

The dependences $PSNR - HVS - M^{nc}(Q)$ and $MS - SSIM^{nc}(Q)$ are monotonous. They are "going jointly" for all test images for $Q < 25$. A joint analysis of all four dependences shows that for $Q < 29$ the introduced distortions are invisible. This means that it is possible to carry out a visually lossless compression and then, after image transferring (storage) and decompression, to perform noise post-filtering if necessary.

In the case of simulations, i.e., if one has a true image, adds noise to it, and then compresses this image, it is also possible to calculate the metric $Metr^{tc}$ between the compressed and true images, and to get the dependence $Metr^{tc}(Q)$. The properties of such dependences are the most interesting. Two examples are given in Figure 3. For $Q < 29$, the image quality is "stable" and it is practically the same as for the uncompressed image. Then, with further Q increasing, three options are possible: (1) the image quality starts to improve and, after attaining the OOP (associated with the dependence maximum), the quality starts to quickly decrease; (2) the quality continues to be almost the same and even local maxima are possible, and then a fast reduction takes the place; (3) the quality starts to decrease. The first situation takes place for the metric $MS - SSIM^{tc}$ for all five images except the test image, Diego (Figure 3b), and for the metric $PSNR - HVS - M^{tc}$ for the test image, Frisco (Figure 3a). The second situation is observed with the metric $PSNR - HVS - M^{tc}$ for some middle complexity test images (Figure 3a). The third situation takes place for the complex structure image, Diego, according to both visual quality metrics (Figure 3a,b). Obviously, the compression in the OOP can be recommended for the first situation, lossy compression in the neighborhood of the local maxima is reasonable in the second situation, and it is unclear what to do in the third situation.

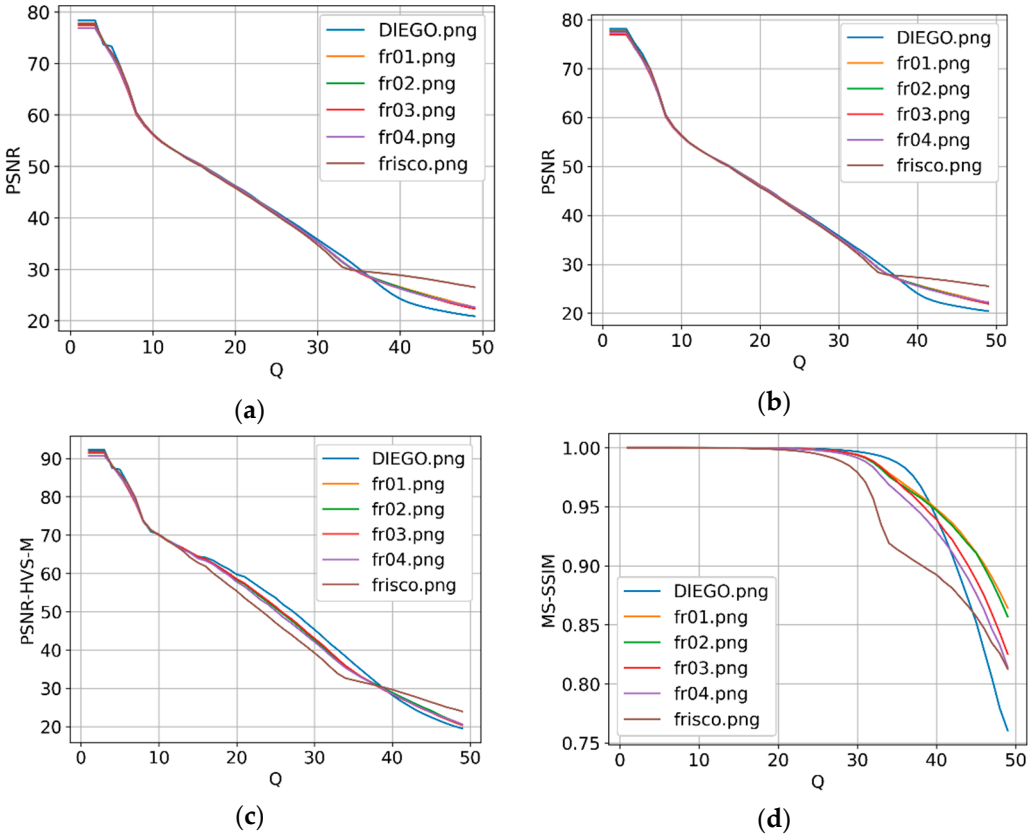


Figure 2. Dependences of the considered metrics (calculated between the original and compressed images) on Q: $PSNR^{nc}(Q)$ for noise variance equal to 64 (**a**), $PSNR^{nc}(Q)$ for noise variance equal to 100 (**b**), $PSNR-HVS-M^{nc}(Q)$ (**c**) and $MS-SSIM^{nc}(Q)$ (**d**) for noise variance equal to 64.

Thus, we come to several questions. The first and the most complicated deals with the fact that in practice one does not have the true image and, thus, the dependence $Metr^{tc}(Q)$ cannot be attained; therefore, the Q in the OOP cannot be determined and it is impossible to understand if the OOP exists or not. Another, less important, problem is what to do if an OOP does not exist.

The arisen question will become clearer after an additional analysis performed below. The plots in Figure 4 visualizes what CR can be provided by a BPG encoder applied to noisy images. Note that the plots in Figure 4 are given for six single-channel RS images of different complexity. The conclusions are the following. First, the CR (for the same Q) depends on the image complexity. For example, for Q = 28 and a noise variance equal to 64, the attained CR for the simplest and most complex test images, Frisco and Diego, differ sufficiently (they are about 4.3 and 3.7, respectively, i.e., one deals with a near-lossless compression). For Q = 35, the situation is the following: the CR for Frisco is about 35.3 whilst the CR for Diego is about 6.8, i.e., the CR differs from a near-lossless case and it is considerably larger for the simple structure image. Second, from comparing the data in Figure 4a,b it is possible to state that the CR for noisier images is smaller. For example, for $\sigma^2 = 196$ and Q = 35, the CR values for the images Frisco and Diego are equal to 5.5 and 6.4, respectively. A sharp increase in the CR starts when Q occurs larger than Q_{OOP}, especially for simple structure images.

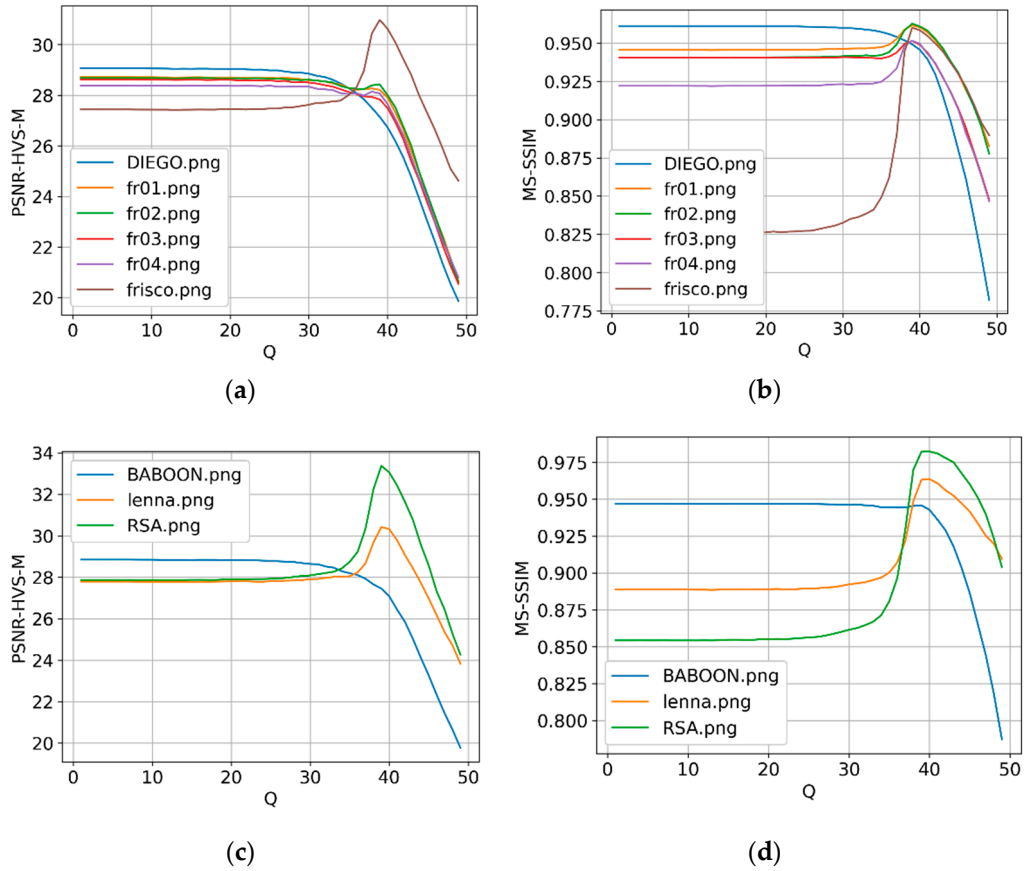


Figure 3. Dependences of the considered metrics (calculated between the true and compressed images on Q: $PSNR-HVS-M^{tc}(Q)$ (**a**,**c**) and $MS-SSIM^{tc}(Q)$ (**b**,**d**) for noise variance equal to 196.

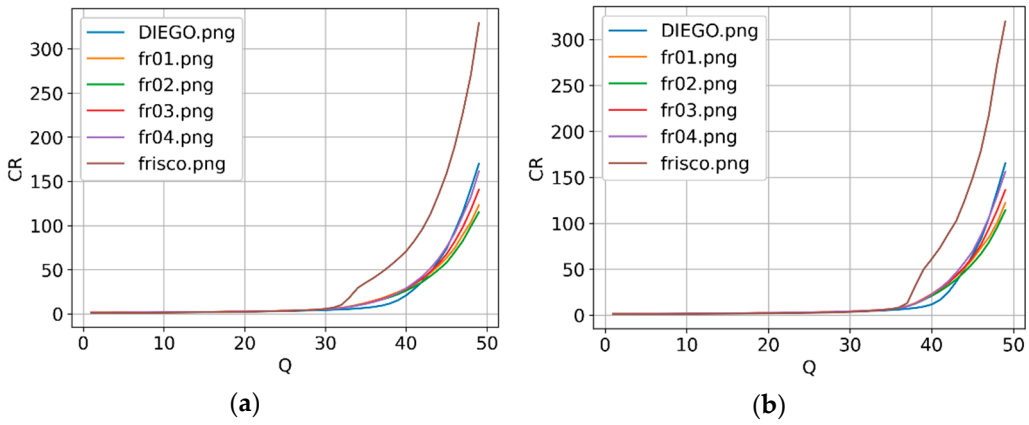
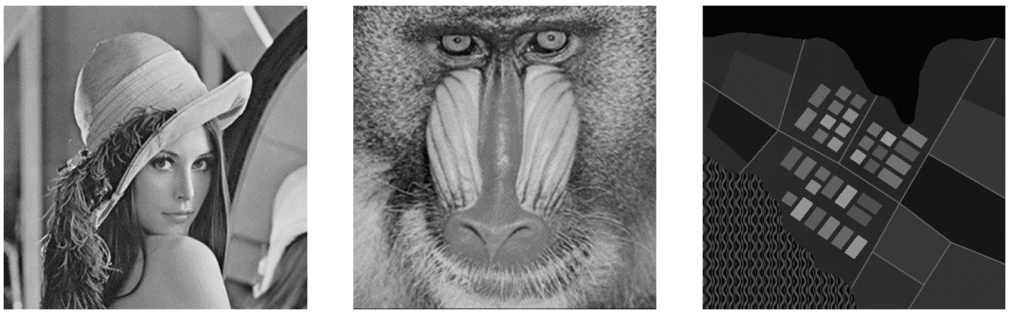


Figure 4. Dependences of CR on Q for noise variance equal to 64 (**a**) and 196 (**b**).

Third, there is no sufficient difference in the general tendencies for images of different origin. To partly prove this, we have obtained the dependences of the considered visual quality metrics on Q for two well-known test images, Lenna and Baboon, as well as an artificial image, RSA (the images are presented in Figure 5). These dependences are given in Figure 3c,d. As one can see, the results for the highly textural images, Diego and Baboon, are very close. Similarly, the results for the simple structure images, RSA, Frisco and Lena are similar as well.



(**a**) (**b**) (**c**)

Figure 5. Test images Lenna (**a**), Baboon (**b**), and RSA (**c**).

3. Properties of Optimal Operation Point

We have already mentioned that we focused on applying the BPG encoder [35]. It has several obvious advantages stimulating to consider its application for RS and other types of images. First, BPG provides a higher compression ratio compared to JPEG and many other encoders for the same quality. Second, BPG is supported by most web browsers. Third, it supports the same formats as JPEG (grayscale, YCbCr 4:2:0, 4:2:2, and 4:4:4), and the most popular RGB, YCgCo and CMYK color spaces are supported. The available versions are able to work with data from 8 to 14 bits per channel. In this paper, we present the results obtained using the grayscale and color (4:2:2) BPG version 0.9.8 given at https://bellard.org/bpg/, accessed on 25 July 2022.

Our main intention in this section is to understand what the Q_{OOP} is and how it can be determined for a given image, metric, and noise variance. The preliminary observations that follow from the plots in Figure 3 are the following. For all images for which OOPs are observed, they take place for approximately the same Q. The OOPs for the metrics, PSNR-HVS-M and MS-SSIM, are observed for practically the same Q. This is a good property that allows to assume that an OOP does not depend on the image at hand and metric used, but, probably, depends on the noise variance.

This hypothesis is based on the results obtained earlier for other coders [24,43–45]. In [43], it has been demonstrated that for an OOP (according to $PSNR^{tc}$) the following condition is satisfied:

$$PSNR^{nc}(Q) \approx PSNR^{n}. \quad (5)$$

This means that for $\sigma^2 = 64$ $PSNR^{nc}(Q) \approx 30$ dB, for $\sigma^2 = 100$ $PSNR^{nc}(Q) \approx 28$ dB and for $\sigma^2 = 196$ $PSNR^{nc}(Q) \approx 25$ dB. Taking into account the expression (4), this should occur for $Q = 33$, 35, and 38, respectively. The plots presented in Figure 6 for noise variances equal to 64 and 100 show that this really is the case. Additional studies carried out for other variance values (25, 144, 196, and 289) and other test images have shown that expression (5) is really valid for $Q = Q_{OOP}$. Substituting (4) into (5) and carrying out simple transformations, it is also possible to obtain (for 8-bit images):

$$Q_{OOP} \approx 14.9 + 20log_{10}(\sigma). \quad (6)$$

Then, knowing the noise variance a priori or estimating it with appropriate accuracy, it is possible to predict for what value of Q the OOP is possible. Meanwhile, the values of $PSNR^{tc}$ for Q calculated according to (6) can differ significantly. For example, for data in Figure 6a, $PSNR^{tc}$ is from 28.5 dB to 36.5 dB where the first case (image Diego) corresponds to an OOP absence whilst the second case (image Frisco) relates to an "obvious" OOP.

The following has been demonstrated in [45] for DCT-based coders. First, OOPs according to visual quality metrics are observed less often than OOPs according to PSNR. Second, an OOP according to visual quality metrics is observed for slightly (by about 5...10%) smaller quantization steps than for an optimal PSNR.

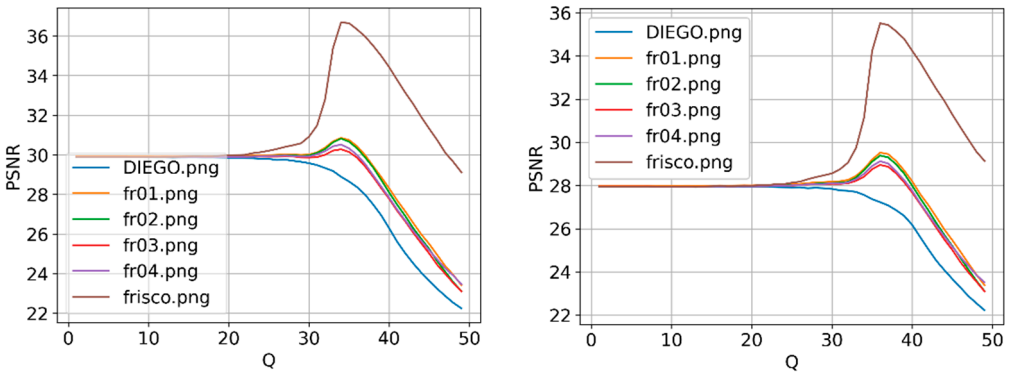


Figure 6. Dependences $PSNR^{tc}(Q)$ for noise variance equal to 64 (**a**) and 100 (**b**).

The same is observed for the BPG encoder. Figure 7 shows dependences $PSNR-HVS-M^{tc}(Q)$ and $MS-SSIM^{tc}(Q)$ for a noise variance equal to 100. They can be compared to the plots in Figure 6b. As one can see, the OOP (if it is observed) takes place for the same Q. Less OOPs are observed according to the visual quality metric $PSNR-HVS-M^{tc}$ (Figure 7a), than according to the standard metric $PSNR^{tc}$ (Figure 6b). Similar effects have been observed for other test images and noise variances.

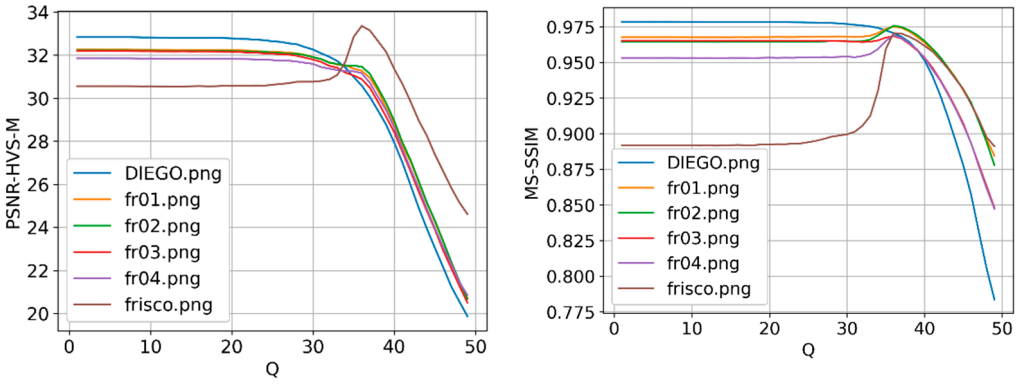


Figure 7. Dependences $PSNR-HVS-M^{tc}(Q)$ (**a**) and $MS-SSIM^{tc}(Q)$ (**b**) for noise variance equal to 100.

Thus, we can suppose that OOPs according to all considered metrics take place for the same Q determined according to (6); however, it might be so that, for a given image and noise variance, an OOP exists according to $PSNR^{tc}$ but does not exist according to $PSNR-HVS-M^{tc}$. Even if the image is compressed in an OOP and the noise is partly removed, residual distortions can be clearly visible ($PSNR-HVS-M^{tc}$ is smaller than 41 dB and $MS-SSIM^{tc}$ is smaller than 0.99).

The carried-out analysis also shows that it is worth analyzing many test images and many values of noise variance to obtain statistical data that allow for making certain conclusions.

4. Prediction of OOP Existence and the Parameters in It

4.1. The Main Idea and Preliminary Results

Our idea of OOP prediction is based on several assumptions. First, we assume that we can predict (estimate) the difference $\Delta Metr = Metr^{tc}(Q_{OOP}) - Metr^{n}$. If this difference is positive, the OOP exists; if negative but quite close to zero, then we deal with situation 2 described in Section 2 (see the examples for four test images in Figure 6a); if negative and its absolute value is large enough, then, being compressed with Q_{OOP} (6), such an image

can be sufficiently degraded. Second, we suppose that such a prediction can be accurate enough to undertake reliable decisions (to avoid making wrong decisions). Third, it is assumed that such a prediction can be performed easily and quickly, i.e., considerably faster than the compression itself.

In fact, we already have experience in solving the aforementioned task—the procedure for predicting the OOP's existence and metric values in the OOP have been proposed in [24]. More in detail, two statistical parameters, $P_{2\sigma}$ and $P_{2.7\sigma}$, that can be easily computed in a set of 8×8 pixel blocks in the DCT domain were considered. For both of them, the expressions that allow for calculating $\Delta PSNR$ and $\Delta PSNR - HVS - M$ using $P_{2\sigma}$ or $P_{2.7\sigma}$ as the argument were obtained. In this paper, we employ the ideas of [24] for the BPG coder, keeping in mind that it is based on DCT and, in this sense, statistics in the DCT domain can be highly correlated with the BPG performance characteristics.

The parameter $P_{2\sigma}$ is calculated as:

$$\begin{gathered} P_{2\sigma} = \sum_{m=1}^{M} \frac{P_{2\sigma}(m)}{M}, \;\; P_{2\sigma}(m) = (\sum_{k=0}^{7} \sum_{l=0}^{7} \delta(k,l,m)\}, \\ \delta(k,l,m) = 1, \; if \; |D(k,l,m)| < 2\sigma \\ 0 \text{ otherwise}, \end{gathered} \tag{7}$$

where M is the number of the considered blocks, $D(k,l,m)$ is the kl-th DCT coefficient in the m-th block, $m = 1, \ldots, M$. In other words, $P_{2\sigma}$ is an estimate of the probability that absolute values of DCT coefficients in blocks are smaller than 2σ supposedly a priori known.

Similarly:

$$\begin{gathered} P_{2.7\sigma} = \sum_{m=1}^{M} \frac{P_{2.7\sigma}(m)}{M}, \;\; P_{2.7\sigma}(m) = (\sum_{k=0}^{7} \sum_{l=0}^{7} \delta(k,l,m)) - 1, \\ \delta(k,l,m) = 1, \; if \; |D(k,l,m)| > 2.7\sigma \\ 0 \text{ otherwise}, \end{gathered} \tag{8}$$

Both statistical parameters have come from the theory of DCT-based denoising [42]. It has been shown there that such parameters are able to jointly characterize the image complexity and noise intensity. For example, $P_{2\sigma}$ is small for complex structure images and/or low intensity noise.

The next task is to obtain dependences that connect $\Delta PSNR$ and $\Delta PSNR - HVS - M$ with the input parameters, $P_{2\sigma}$ or $P_{2.7\sigma}$. This stage is carried out off-line and can be treated as specific training. The main task is, such as in machine learning, to take into account all possible situations, i.e., a wide variety of image properties and noise intensities.

At this stage, scatter-plots of $\Delta Metr$ on the input parameters have been formed. Each scatter-plot has been obtained as follows: noise with a given noise variance has been added to a considered test image that has then been compressed using Q_{OOP} (6). Then, the input and output parameters have been determined. In general, eleven test images of different complexity have been exploited and the noise variance was varied in the limits from 0.25 to 400. The obtained scatter-plots are presented in Figures 8–11.

Preliminary analysis shows the following:

1. The points of the scatter-plots $\Delta PSNR$ vs. $P_{2\sigma}$ (Figure 8) and $\Delta PSNR$ vs. $P_{2.7\sigma}$ (Figure 10) are placed in a very compact manner, clearly showing that there are monotonous increasing and decreasing dependences in the output parameter on the input parameter, respectively; both the input parameters are in the limits from zero to unity, for small $P_{2\sigma}$ and large $P_{2.7\sigma}$, $\Delta PSNR$ are negative and close to -2 dB (as an example, see the data for the test image Diego in Figure 5a), and in this case the OOP is absent and it is not worth compressing the images using Q_{OOP} (6);
2. The points of the scatter-plots $\Delta PSNR - HVS - M$ vs. $P_{2\sigma}$ (Figure 9) and $\Delta PSNR - HVS - M$ vs. $P_{2.7\sigma}$ (Figure 11) are placed in a less compact manner; meanwhile, there are obvious tendencies of $\Delta PSNR - HVS - M$ increasing with $P_{2\sigma}$ increasing

and $P_{2.7\sigma}$ decreasing; a visual analysis allows supposing that these dependencies are monotonous; the scatter-plot points are placed more compactly for large $P_{2\sigma}$ and small $P_{2.7\sigma}$ that correspond to simpler structure images and/or more intensive noise;

3. It is possible to expect that a good curve fitting is possible in both cases that allows establishing approximate analytic dependences between the output and input parameters (examples of the fitted curves are given in all Figures 8–11);
4. There are points where these curves cross the zero level: $P_{2\sigma} \approx 0.73$ and $P_{2.7\sigma} \approx 0.16$ for $\Delta PSNR$; $P_{2\sigma} \approx 0.84$ and $P_{2.7\sigma} \approx 0.05$ for $\Delta PSNR - HVS - M$.

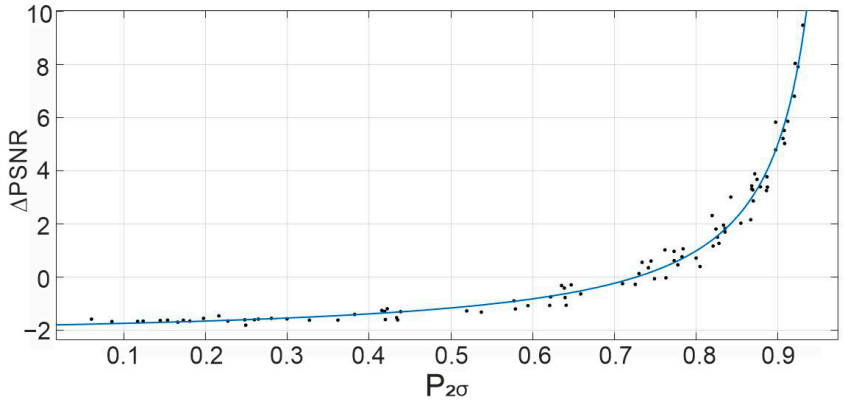


Figure 8. The scatter-plot $\Delta PSNR$ vs. $P_{2\sigma}$ and the fitted curve.

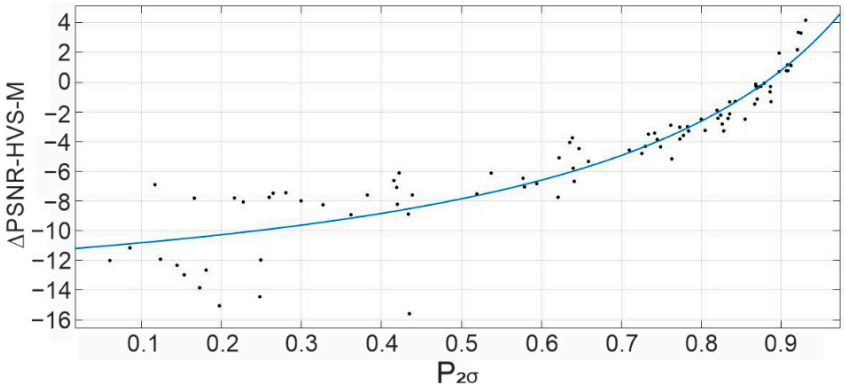


Figure 9. The scatter-plot $\Delta PSNR - HVS - M$ vs. $P_{2\sigma}$ and the fitted curve.

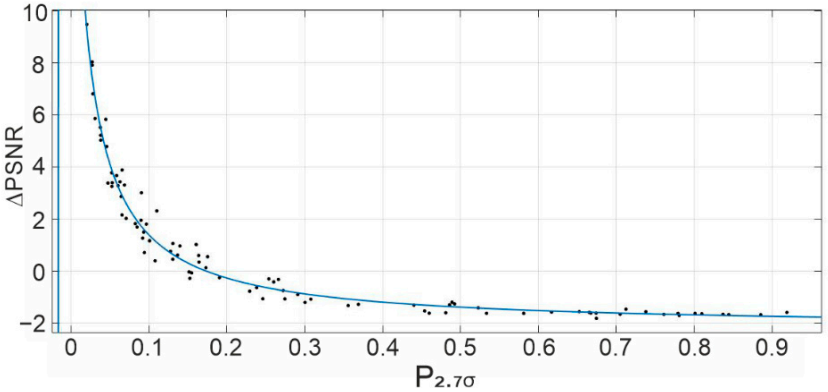


Figure 10. The scatter-plot $\Delta PSNR$ vs. $P_{2.7\sigma}$ and the fitted curve.

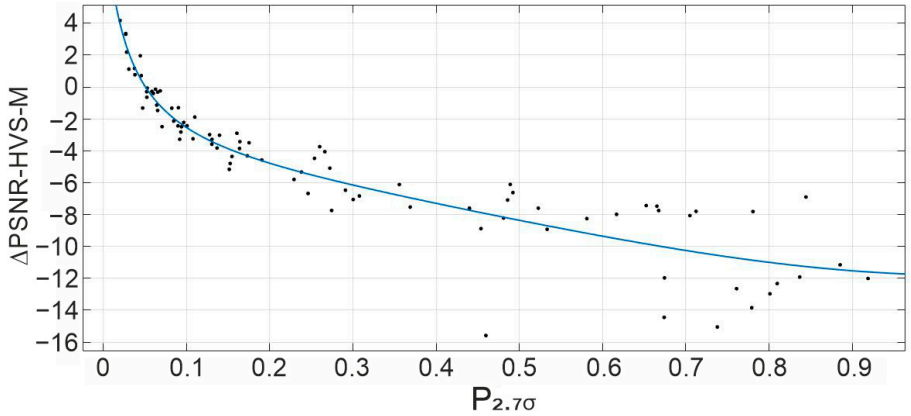


Figure 11. The scatter-plot $\Delta PSNR - HVS - M$ vs. $P_{2.7\sigma}$ and the fitted curve.

4.2. Curve Fitting Details

Having, in general, shown the possibility of a good curve fitting, we come to the next questions—what the best (appropriate) fitting is and how to characterize its accuracy. There are well developed theories of LMSE and robust fitting [46,47]. A visual analysis of the scatter-plots in Figures 7–10 shows that there are no obvious outliers in the data; thus, there is no obvious necessity to apply a robust fitting. LMSE regression is implemented in many software tools including MATLAB and others; therefore, we can use one of them (the MATLAB Curve Fitting Tool in our case). To characterize the fitting accuracy, it is common [46] to use the root mean square (RMSE) that should be as small as possible, as well as a goodness-of-the-fit R^2 and/or adjusted goodness-of-the-fit $AdjR^2$ that have to be as large (close to the unity) as possible.

Standard curve fitting tools usually provide several options of fitting functions including polynomials, exponential or weighted sum of exponents, rational functions, Fourier series, etc. Finding the best fitting function is more an engineering and heuristic rather than scientific task. Moreover, some solutions can be very close to each other according to quantitative criteria that does not allow for giving a unique practical recommendation.

The curves fitted for the scatter-plots in Figures 8–11 have been obtained using rational functions. For the data in Figure 8, one has an RMSE about 0.37 and R^2 and $AdjR^2$ about 0.98. For the scatter-plot in Figure 9, the RMSE is about 1.72 and the R^2 and $AdjR^2$ are about 0.85. For the data in Figure 10, one has an RMSE about 0.43 and R^2 and $AdjR^2$ about 0.975. For the scatter-plot in Figure 11, the RMSE is about 1.66 and the R^2 and $AdjR^2$ are about 0.86. Supposing that the fitting is good enough, it is possible to conclude that the fitting for $\Delta PSNR$ is considerably more accurate than for $\Delta PSNR - HVS - M$ (this is not because of bad fitting but because of larger sparsity of data for $\Delta PSNR - HVS - M$). It is better to use $P_{2\sigma}$ in predicting the $\Delta PSNR$ and $P_{2.7\sigma}$ in predicting the $\Delta PSNR - HVS - M$. The parameters of the fitted curves are presented in Table 1.

One might be interested whether or not other approximations are able to provide better results. The Fourier models, e.g., the Fourier model 3, are able to provide quite a good fitting in the sense of quantitative criteria (see Figure 12), but the obtained curves can be not monotonous, and that is not in agreement with our assumptions. The same relates to polynomial approximations. For polynomials of the order 2 and 3, the fitting accuracy parameters are worse than reported above, whilst for polynomials for the order 4 and 5, the curves become not monotonous. Finally, good approximations can be obtained for the weighted sums of two exponentials (see Figure 13 and compare it to Figure 9); thus, the use of rational functions or the sums of two exponentials can be considered an appropriate choice.

Table 1. Parameters of the fitted curves.

Dependence	Expression	Parameters
$\Delta PSNR$ on $P_{2\sigma}$	$f(x) = (p_1x + p_2)/(x^3 + q_1x^2 + q_2x + q_3)$	$p_1 = 1.533 \times 10^4$, $p_2 = -1.112 \times 10^4$, $q_1 = 75.71$, $q_2 = -6291$, $q_3 = 6139$
$\Delta PSNR$ on $P_{2.7\sigma}$	$f(x) = (p_1x + p_2)/(x^2 + q_1{*}x + q_2)$	$p_1 = -6.162 \times 10^5$, $p_2 = 1.077 \times 10^5$, $q_1 = 2.843 \times 10^5$, $q_2 = 4501$
$\Delta PSNR - HVS - M$ on $P_{2\sigma}$	$f(x) = (p_1x + p_2)/(x^2 + q_1x + q_2)$	$p_1 = 2.895 \times 10^5$, $p_2 = -2.549 \times 10^5$, $q_1 = -1.72 \times 10^4$, $q_2 = 2.263 \times 10^4$
$\Delta PSNR - HVS - M$ on $P_{2.7\sigma}$	$f(x) = (p_1x + p_2)/(x^3 + q_1x^2 + q_2x + q_3)$	$p_1 = -10.97$, $p_2 = 0.558$, $q_1 = -1.99$, $q_2 = 1.82$, $q_3 = 0.048$

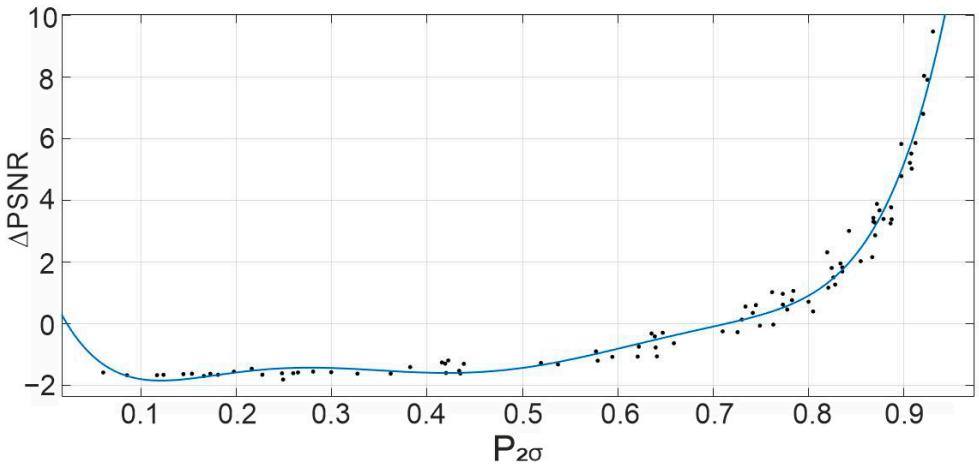


Figure 12. The scatter-plot $\Delta PSNR$ vs. $P_{2\sigma}$ and the fitted curve of the model Fourier 3 (RMSE is about 0.42 and the R^2 and AdjR^2 are about 0.976).

Having the fitted curves at hand, it is possible to predict $\Delta PSNR$ and $\Delta PSNR - HVS - M$ for any image to be compressed. For this purpose, one has to calculate the input parameter and to substitute it into the approximating curve defined in Table 1. For example, the estimated $P_{2\sigma}$ is equal to 0.6. Then, according to the data in Figure 7 (or Figure 12), the predicted $\Delta PSNR$ is about -0.8 dB, and the predicted $\Delta PSNR - HVS - M$ is about -6.5 dB according to the data in Figure 8 and about -6.1 dB according to the approximation in Figure 13. In any case, one can conclude that the OOP for this image is absent and that it is not worth compressing it with Q_{OOP}. An example of such a situation is given in Figure 14, where both the predicted and the true $\Delta PSNR$ and $\Delta PSNR - HVS - M$ are negative (the predicted values are equal to -0.71 dB and -6.34 dB whilst the true values are equal to -0.73 dB and -5.09 dB, respectively). As one can see, the compressed image is slightly smeared.

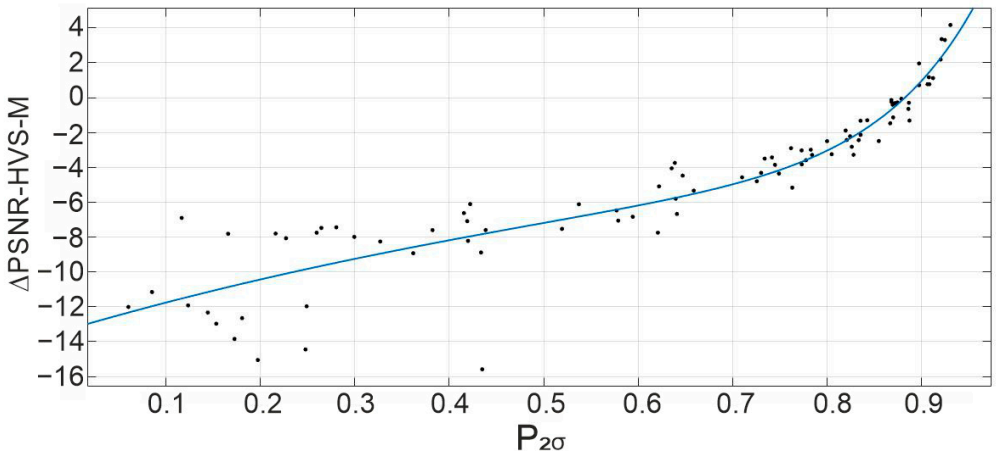


Figure 13. The scatter-plot $\Delta PSNR - HVS - M$ vs. $P_{2\sigma}$ and the fitted curve (the weighted sum of two exponentials, RMSE is about 1.64 and the R^2 and AdjR^2 are about 0.87).



(**a**) (**b**) (**c**)

Figure 14. The noise-free test image Fr03 (**a**), this image corrupted by AWGN with noise variance equal to 25 (**b**) and the image compressed using Q_{OOP} = 29 (**c**), CR = 5.84.

One can argue that the prediction is not accurate for $\Delta PSNR - HVS - M$ when $P_{2\sigma}$ is about 0.2. Consequently, in this case, $\Delta PSNR - HVS - M$ can be from -16 dB to -7 dB depending on the image at hand whilst the predicted values are about -10.5 dB (see Figure 13). Such an accuracy of prediction seems low since the errors can be up to 5.5 dB; however, in practice, this is not a problem for several reasons. First, such a small $P_{2\sigma}$ usually corresponds to images corrupted by low intensity noise which is invisible. Then, no positive effect of noise filtering can be seen anyway. Second, such $\Delta PSNR - HVS - M$ values usually correspond to $PSNR - HVS - M^n$ about 55...60 dB. Then, after compression, one has a $PSNR - HVS - M^{tc}$ about 45...50 dB, i.e., the introduced distortions are invisible.

Analyzing the scatter-plots for $\Delta PSNR - HVS - M$, one can expect that the prediction accuracy for $P_{2\sigma} > 0.5$, which is of more interest, is considerably better than for $P_{2\sigma} \leq 0.5$ (see, e.g., the scatter-plot in Figure 9). To check this hypothesis, we have calculated the RMSE of the fitted curves in the considered intervals. The RMSE for $P_{2\sigma} \leq 0.5$ equals to 2.34 whilst it is equal to 1.21 for $P_{2\sigma} > 0.5$, i.e., our assumption is valid. Then, one can carry out quite an accurate prediction in the interval $P_{2\sigma} > 0.5$ where it is especially important to undertake decisions on the OOP's existence and Q setting.

4.3. Factors Affecting the Accuracy of Prediction

There are also other factors affecting the accuracy of prediction. A first factor is noise realization. It is clear that, for a given test image and AWGN variance, the $\Delta PSNR$,

$\Delta PSNR - HVS - M$, and even Q_{OOP} can vary from one noise realization to another. To study this aspect, we have analyzed three images of different complexity corrupted by AWGN with three different values. The values of $\Delta PSNR$ and $\Delta PSNR - HVS - M$ have been measured for Q_{OOP} and the variances of these parameters have been calculated. It has been established that variances of these parameters are about 0.001; even in the worst case the maximal variance was equal to 0.0033 for $\Delta PSNR$ and 0.0046 for $\Delta PSNR - HVS - M$. This means that the influence of realization can be neglected (recall that the RMSE of fitting is about 0.4 for $\Delta PSNR$ and about 1.2 for $\Delta PSNR - HVS - M$ in the area under interest). If the Q_{OOP} was different for different realizations, the differences were equal to 1 (e.g., an OOP was observed for Q equal to either 33 or 34) and the difference in values of $\Delta PSNR$ or $\Delta PSNR - HVS - M$ in the neighbor points was negligible (e.g., analyze the dependences in Figure 3 near their peaks). Thus, in practice, the influence of noise realization on the parameters of the OOP can be ignored.

One more parameter that can influence the prediction is a possible variation of the input parameter. Having an approximation $\Delta Metr(Q)$ and variance of the input parameter σ_P^2, a variance of $\Delta Metr(Q)$ can be estimated as $Var^2_{\Delta Metr} \approx \sigma_P^2 \left(\frac{d\Delta Metr(Q)}{dQ}\right)^2$. Then, $\Delta Metr(Q)$ can be large in places where the absolute values of $\frac{d\Delta Metr(Q)}{dQ}$ are large and if σ_P^2 is quite large. The absolute values of $\frac{d\Delta Metr(Q)}{dQ}$ are large for a large $P_{2\sigma}$ or, equivalently, a small $P_{2.7\sigma}$; however, we are more interested in places where the approximation curves cross the zero level. The absolute values of the derivatives there are about 50; thus, we need to understand what the values are of σ_P^2.

It is possible to expect that σ_P^2 might depend on several factors, namely, the number of blocks M, the properties of images and noise variance, and how the blocks are positioned in a considered image. To avoid possible problems with the images with some regular structures, we propose to apply a random positioning of blocks where indices of their left upper corner are (rounded-off) random variables in the limits from $i = 1$ till $i = I_{Im} - 7$ and from $j = 1$ till $j = J_{Im} - 7$, with a uniform distribution. Within this approach, we have first analyzed three test images of different complexity with three different noise variance values. A set of realizations of the random block positions have been generated for $M = 500$. A variance σ_P^2 for the input parameter $P_{2\sigma}$ was in the limits from 5×10^{-6} till 8×10^{-5}. There is no obvious dependence on the image complexity but there is a tendency of σ_P^2 decreasing if the AWGN variance increases (σ_P^2 is about 0.00001 for an AWGN variance about 200 that usually corresponds to $P_{2\sigma} \geq 0.5$). The latter is a positive factor since $\left(\frac{d\Delta Metr(Q)}{dQ}\right)^2$ is larger just for a large $P_{2\sigma}$. Then, $\sigma_P^2 \left(\frac{d\Delta Metr(Q)}{dQ}\right)^2$ is about 2.5×10^{-2} (the standard deviation is about 0.16 and it is sufficiently smaller than the RMSE of prediction due to fitting).

We have also carried out a similar study for the input parameter $P_{2.7\sigma}$. The limits of its variance variation are similar—from 3×10^{-6} till 8×10^{-5}. The σ_P^2 values are smaller for a larger AWGN variance.

In addition, we have also fixed the positions of the blocks and calculated the variances σ_P^2 for a set of noise realizations. This time the variances were even smaller (from 2×10^{-6} till 6×10^{-6}) and, hence, we can neglect this factor.

4.4. Other Practical Aspects

Note that it is quite easy and fast to carry out a prediction. A calculation of the DCT in 8×8 blocks is a standard operation in image processing [48] that can be performed efficiently using both hardware and software. Then, easy comparisons and elementary arithmetic operations are needed to calculate the input and output parameters.

One can also be interested in whether or not the lossy compression by BPG in the OOP produces some benefits compared to other lossy compression techniques applied to noisy images. Table 2 contains data that allow for comparing BPG to the coder AGU [49] for some test images and some noise variances. First, a comparison has been completed for two images used in obtaining the approximating dependences (fitted curves), namely,

the images, Frisco and Fr01. As one can see, the BPG encoder provides benefits in three senses: (1) it produces about a 0.8 dB better $PSNR^{nc}$ in the OOP; (2) it provides a better visual quality according to the visual quality metric $MS-SSIM^{nc}$, and (3) a slightly (by a few percent) larger CR is usually ensured.

Table 2. Obtained simulation data.

Test Image	AWGN Variance	AGU			BPG		
		PSNR (dB)	CR	MS-SSIM	PSNR (dB)	CR	MS-SSIM
Frisco	64	36.05	31.50	0.972	36.83	32.84	0.977
	100	34.84	38.36	0.965	35.68	41.17	0.972
	196	33.09	54.93	0.949	33.83	56.63	0.960
Fr01	100	28.26	9.33	0.966	29.66	10.09	0.976
	196	26.66	14.3	0.950	27.78	14.3	0.963
Aerial	100	28.29	8.87	0.985	29.63	9.54	0.976
	196	26.57	13.03	0.946	27.67	13.04	0.962
Airfield	100	27.30 *	7.69	0.951 *	28.25	10.52	0.956
	196	25.79	12.35	0.927 *	26.69	12.34	0.938

* OOP is absent.

Meanwhile, it can also be interesting to check if this happens for images not used in obtaining the scatter-plots (for image processing approaches based on learning, a verification stage is obligatory). For this purpose, we have obtained simulation data for two test images—Aerial and Airfield—that have not been used in previous analysis. The results are presented in Table 2. Their analysis shows the following. First, again the results for the BPG coder are better for both metrics and, at least, not worse according to the CR. An interesting situation has been observed for the test image, Aerial, corrupted by AWGN with a variance equal to 64 and the image, Airfield, corrupted by AWGN with a noise variance equal to 100. For the encoder AGU, an OOP is absent whilst for the encoder BPG, an OOP exists for both the considered metrics.

Thus, we can state that the BPG encoder outperforms AGU and this takes place for both images that have been used in obtaining scatter-plots and images that have been employed for verification.

5. Decision-Making and Other Practical Cases

Supposing that prediction has been carried out for a given image, one obtains $\Delta PSNR$ or $\Delta PSNR-HVS-M$ or both. A question is what to do further? There can be different options—to rely only on one parameter (rather than one of two) or on both. The choice is very heuristic without an obvious preference (e.g., special studies with customers assessing the compressed image quality can be carried out to obtain a reliable answer). We propose the following practical algorithm:

1. If $\Delta PSNR + \Delta PSNR-HVS-M > 1$ dB, consider that an OOP exists and compress an image using Q_{OOP};
2. If -1 dB $\leq \Delta PSNR + \Delta PSNR-HVS-M \leq 1$ dB, consider that an OOP might exist and use Q_{OOP}-1 (but not less than 28) to avoid oversmoothing;
3. If $\Delta PSNR + \Delta PSNR-HVS-M \leq -1$ dB, use Q = 28 to provide invisibility of introduced distortions and the possibility of noise removal in it by applying post-filtering to the decompressed images.

An example in Figure 13 corresponds to case 3. An example for situation 1 is presented in Figure 15. The positive effect of a lossy compression is seen, especially in an image's homogeneous regions where the noise removal is obvious. Finally, an example for situation 2 is demonstrated in Figure 16. In this case, the positive effect of noise suppression can be noticed as well; however, edge/detail smearing takes place too.

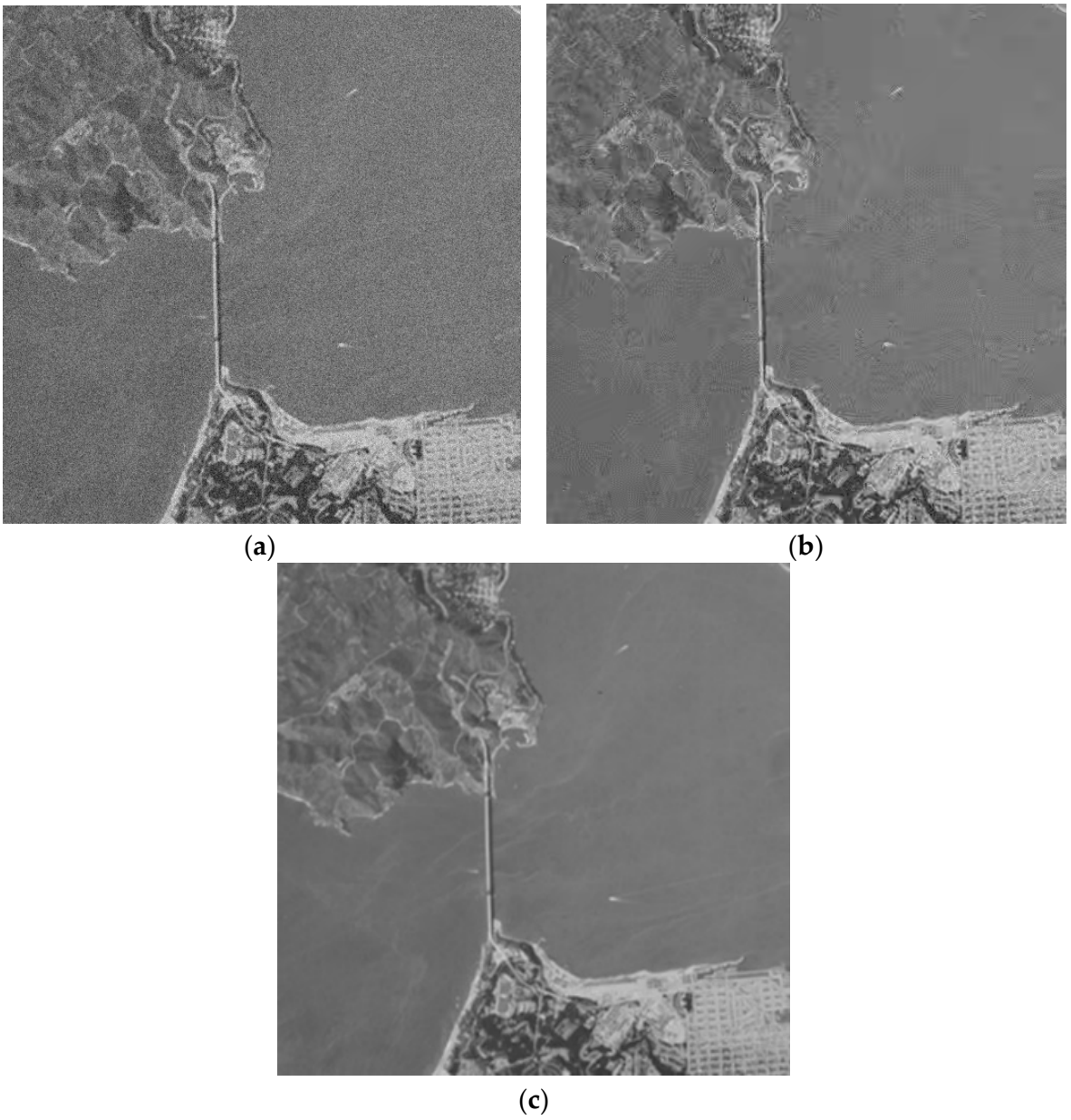


Figure 15. Noisy image Frisco (**a**) and the same image compressed in OOP (**b**), noise variance equal to 196, and the noise-free image (**c**).

If situation 1 is falsely recognized as situation 2 or vice versa, practically nothing happens since the recommended Q values differ only by unity. If situation 2 is falsely recognized as situation 3, it is not a problem since a "careful" compression is carried out (by the expense of a smaller CR). If situation 3 is falsely recognized as situation 2, a slightly more smearing of the compressed image can be observed (however, the CR is larger). Finally, situations 1 and 3 are practically never misclassified.

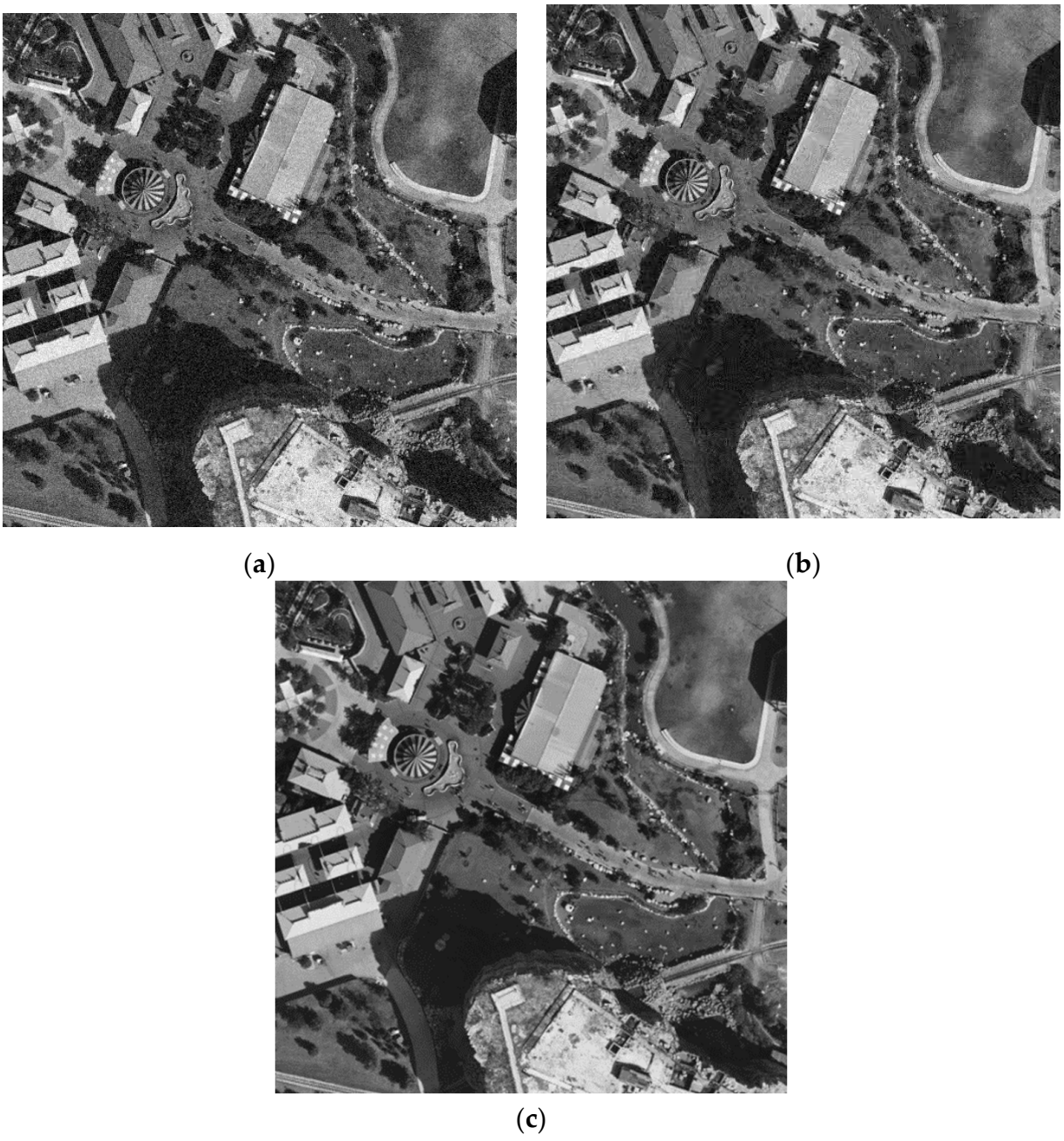


Figure 16. Noisy image Fr02 (**a**) with noise variance equal to 196 and the same image compressed using Q_{OOP} = 37 (**b**), the noise-free image (**c**).

Thus, the fully automatic procedure for a given noisy image is as follows:

1. Estimate the noise variance by some blind method of a noise variance assessment (if the noise variance is not known in advance);
2. Calculate Q_{OOP} according to (6); calculate $P_{2\sigma}$ according to (7);
3. Calculate $\Delta PSNR$ and $\Delta PSNR - HVS - M$ using the expressions and parameters in Table 1;
4. Make a decision on the recommended Q_{OOP} according to the algorithm described above;
5. Carry out a compression using the recommended Q_{OOP}.

As it has been stated in Section 2, AWGN can be a simplified noise model for which the noise variance or standard deviation can be estimated in a blind manner (automatically) [50,51]. A more general case is the signal-dependent noise model [52]. Let us

demonstrate that the OOP is also possible for images corrupted by Poisson noise compressed by BPG. Note that two approaches to such a compression are possible. The first one is to apply BPG directly. The second one is to apply a proper VST (Anscombe transform in the considered case [53]) and some pre-normalization before the compression with inverse operations at the decompression stage. In this case, the signal-dependent noise converts to an almost pure additive and we arrive at the image/noise model studied above; thus, let us concentrate on the first approach.

The analysis has been carried out for six RS test images. The obtained rate-distortion curves $PSNR-HVS-M^{tc}(Q)$ and $MS-SSIM^{tc}(Q)$ are given in Figure 17. As one can see, the OOP exists for the metric MS-SSIM for the images, Frisco and Fr04, and for the metric PSNR-HVS-M for the image, Frisco (similar situations have been earlier observed for the AWGN case, see the plots for the test image Fr02 in Figure 7). In all three cases, the OOP is observed for $Q = 37$. According to $PSNR^{tc}$, the OOP exists as well and this happens for $Q = 37$, while $\Delta PSNR$ can reach 3 dB; thus, we can expect that automatic procedures for lossy compression can be designed for BPG for images corrupted by signal-dependent noise, including the speckle typical for synthetic aperture radar images.

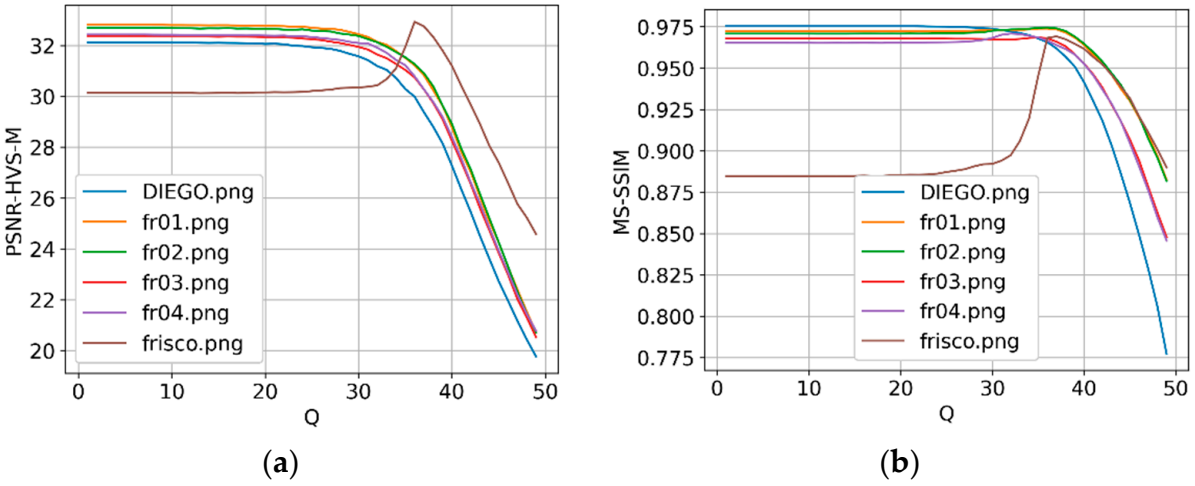


Figure 17. Dependences $PSNR-HVS-M^{tc}(Q)$ (**a**) and $MS-SSIM^{tc}(Q)$ (**b**) for six test images corrupted by Poisson noise.

It might also be true that two, three or more image components are corrupted by the noise and these images have to be compressed in a lossy manner. Certainly, a component-wise compression for which all the results obtained above are valid is possible; however, the joint compression of several component noisy images is possible as well. For example, available BPG software allows for compressing color, i.e., three-channel images; thus, it is easily possible to test this practical situation. As an initial case, let us assume that all the component images are corrupted by AWGN with the same noise variance. Since the results can be of interest for both three-channel and color images, four test images have been considered: the RS three-channel images, Frisco and Diego, and the widely used color test images, Lena and Baboon. The original images of size 512×512 pixels were presented in RGB, an AWGN with variance equal to 100 was independently added to each component image, a 4:2:2 version of BPG was applied, and the metrics were calculated independently for each component image. The obtained dependences are presented in Figure 18.

It follows from the preliminary analysis that the OOP can also exist and it takes place practically for the same Q for all component images. Meanwhile, there are also some specific results, such as more obvious OOPs for the G component; thus, additional studies are needed.

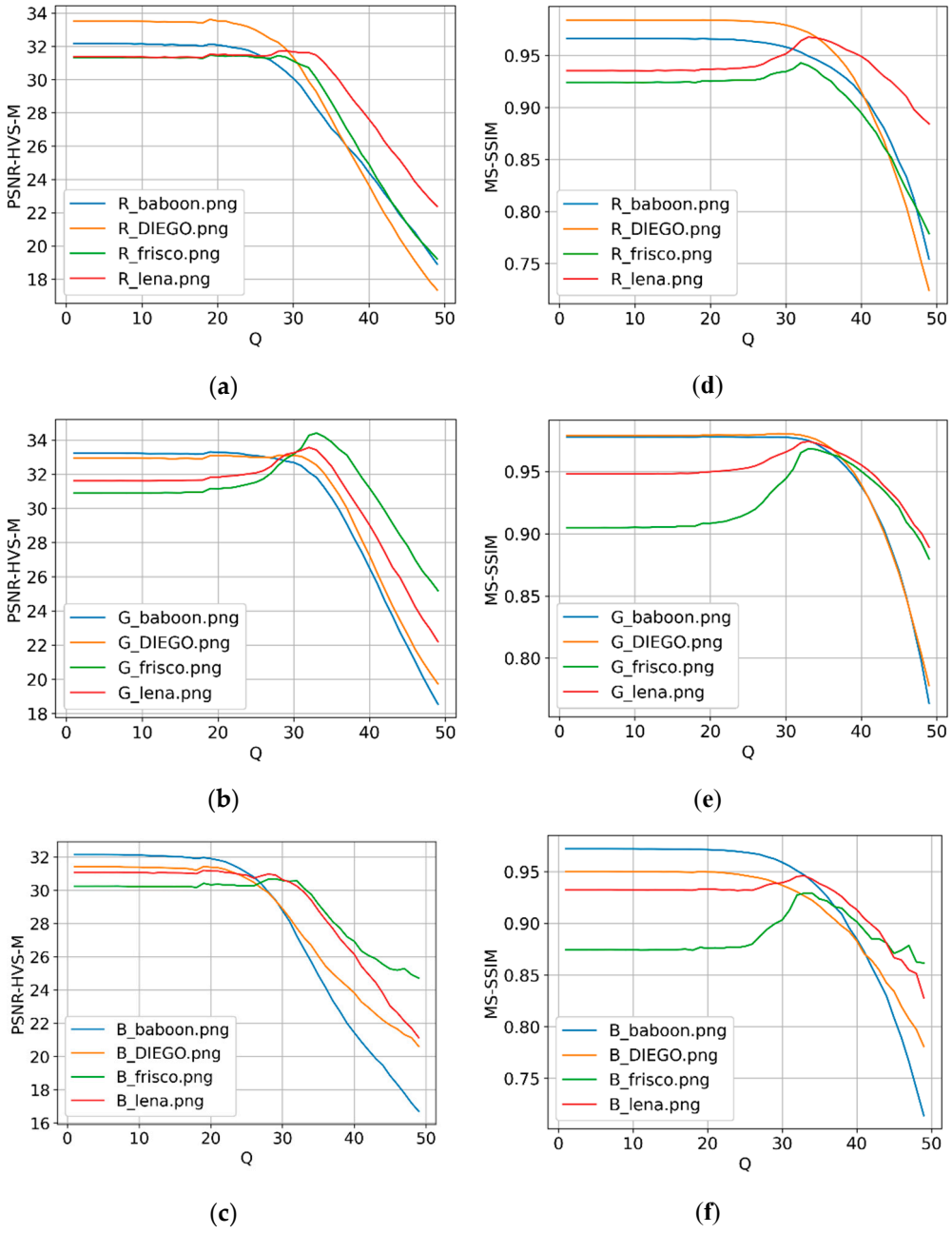


Figure 18. Dependences $PSNR - HVS - M^{tc}(Q)$ for R (**a**), G (**b**), and B (**c**), and $MS - SSIM^{tc}(Q)$ dependences for R (**d**), G (**e**), and B (**f**) component images.

6. Conclusions

The task of the lossy compression of images corrupted by AWGN by the BPG coder is considered. The main attention is paid to the single-channel image case. It is shown that an OOP might exist according to the standard criteria as PSNR (and MSE). Moreover, an OOP might exist for the visual quality metrics such as MS-SSIM and PSNR-HVS-M, although, this occurs more rarely. It is demonstrated that OOP existence depends on the noise variance and image complexity where an OOP exists with a higher probability for simpler structure images and/or a higher intensity noise. With a noise intensity increase, the Q_{OOP} increases as well, and the corresponding expression is obtained. The approach to predicting the OOP's existence and the metric values in it is proposed. This is based on obtaining the approximating dependences in advance using scatter-plots and curve fitting. Then, for a given image, a simple statistical parameter has to be estimated, its value

has to be used as an approximator input and the prediction with decision-making needs to be completed. Its accuracy is analyzed and shown to be appropriate for practice. The recommendations on an automatic Q setting for different practical situations are given and illustrated.

The possibility of OOP existence for signal-dependent noise and three-channel images corrupted by AWGN is shown. This can be the direction of the future research. Another direction deals with improving the prediction accuracy for visual quality metrics by the joint processing of several input parameters that can be performed by a trained neural network.

Author Contributions: Conceptualization, V.L. and B.V.; methodology, S.K.; software, B.K.; validation, V.N., B.K. and S.K.; formal analysis, B.K. and V.L.; investigation, B.K.; writing—original draft preparation, V.L.; writing—review and editing, B.V.; visualization, B.K., S.K. and V.N.; supervision, V.L. and B.V. All authors have read and agreed to the published version of the manuscript.

Funding: The research performed in this manuscript was partially supported by the French Ministries of Europe and Foreign Affairs (MEAE) and Higher Education, Research and Innovation (MESRI) through the PHC Dnipro 2021 project No. 46844Z.

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mielke, C.; Boesche, N.K.; Rogass, C.; Segl, K.; Gauert, C.; Kaufmann, H. Potential Applications of the Sentinel-2 Multispectral Sensor and the Enmap Hyperspectral Sensor in Mineral Exploration. *EARSeL Eproceedings* **2014**, *13*, 93.
2. Schowengerdt, R.A. *Remote Sensing, Models, and Methods for Image Processing*, 3rd ed.; Academic Press: Burlington, MA, USA, 2007; ISBN 978-0-12-369407-2.
3. Kussul, N.; Lavreniuk, M.; Kolotii, A.; Skakun, S.; Rakoid, O.; Shumilo, L. A Workflow for Sustainable Development Goals Indicators Assessment Based on High-Resolution Satellite Data. *Int. J. Digit. Earth* **2020**, *13*, 309–321. [CrossRef]
4. Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.; et al. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sens.* **2016**, *8*, 70. [CrossRef]
5. Khorram, S.; van der Wiele, C.F.; Koch, F.H.; Nelson, S.A.C.; Potts, M.D. Future Trends in Remote Sensing. In *Principles of Applied Remote Sensing*; Springer International Publishing: Cham, Switzerland, 2016; pp. 277–285, ISBN 978-3-319-22559-3.
6. Swarnalatha, P.; Sevugan, P. (Eds.) *Big Data Analytics for Satellite Image Processing and Remote Sensing*; Advances in Computer and Electrical Engineering; IGI Global: Philadelphia, PA, USA, 2018; ISBN 978-1-5225-3643-7.
7. Blanes, I.; Magli, E.; Serra-Sagrista, J. A Tutorial on Image Compression for Optical Space Imaging Systems. *IEEE Geosci. Remote Sens. Mag.* **2014**, *2*, 8–26. [CrossRef]
8. Chow, K.; Tzamarias, D.; Blanes, I.; Serra-Sagristà, J. Using Predictive and Differential Methods with K2-Raster Compact Data Structure for Hyperspectral Image Lossless Compression. *Remote Sens.* **2019**, *11*, 2461. [CrossRef]
9. Radosavljević, M.; Brkljač, B.; Lugonja, P.; Crnojević, V.; Trpovski, Ž.; Xiong, Z.; Vukobratović, D. Lossy Compression of Multispectral Satellite Images with Application to Crop Thematic Mapping: A HEVC Comparative Study. *Remote Sens.* **2020**, *12*, 1590. [CrossRef]
10. Aiazzi, B.; Alparone, L.; Baronti, S. Near-Lossless Compression of 3-D Optical Data. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2547–2557. [CrossRef]
11. Santos, L.; Lopez, S.; Callico, G.M.; Lopez, J.F.; Sarmiento, R. Performance Evaluation of the H.264/AVC Video Coding Standard for Lossy Hyperspectral Image Compression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 451–461. [CrossRef]
12. Penna, B.; Tillo, T.; Magli, E.; Olmo, G. Transform Coding Techniques for Lossy Hyperspectral Data Compression. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1408–1421. [CrossRef]
13. Christophe, E. Hyperspectral Data Compression Tradeoff. In *Optical Remote Sensing*; Prasad, S., Bruce, L.M., Chanussot, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 9–29, ISBN 978-3-642-14211-6.
14. Yu, G.; Vladimirova, T.; Sweeting, M.N. Image Compression Systems on Board Satellites. *Acta Astronaut.* **2009**, *64*, 988–1005. [CrossRef]
15. Ozah, N.; Kolokolova, A. Compression Improves Image Classification Accuracy. In *Advances in Artificial Intelligence*; Meurs, M.-J., Rudzicz, F., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; Volume 11489, pp. 525–530, ISBN 978-3-030-18304-2.
16. Chen, Z.; Hu, Y.; Zhang, Y. Effects of Compression on Remote Sensing Image Classification Based on Fractal Analysis. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4577–4590. [CrossRef]

17. Vasilyeva, I.; Li, F.; Abramov, S.K.; Lukin, V.V.; Vozel, B.; Chehdi, K. Lossy Compression of Three-Channel Remote Sensing Images with Controllable Quality. In Proceedings of the Image and Signal Processing for Remote Sensing XXVII, Online, Spain, 13–17 September 2021; Bruzzone, L., Bovolo, F., Benediktsson, J.A., Eds.; SPIE: Madrid, Spain, 2021; p. 26.
18. Lee, J.-S.; Pottier, E. *Polarimetric Radar Imaging: From Basics to Applications*; Optical science and engineering; CRC Press: Boca Raton, FL, USA, 2009; ISBN 978-1-4200-5497-2.
19. Mullissa, A.G.; Persello, C.; Tolpekin, V. Fully Convolutional Networks for Multi-Temporal SAR Image Classification. In Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 23–27 July 2018; IEEE: Valencia, Spain, 2018; pp. 6635–6638.
20. Zhong, P.; Wang, R. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2260–2275. [CrossRef]
21. Wang, W.; Zhong, X.; Su, Z. On-Orbit Signal-to-Noise Ratio Test Method for Night-Light Camera in Luojia 1-01 Satellite Based on Time-Sequence Imagery. *Sensors* **2019**, *19*, 4077. [CrossRef] [PubMed]
22. Al-Shaykh, O.K.; Mersereau, R.M. Lossy Compression of Noisy Images. *IEEE Trans. Image Process.* **1998**, *7*, 1641–1652. [CrossRef]
23. Chang, S.G.; Yu, B.; Vetterli, M. Image Denoising via Lossy Compression and Wavelet Thresholding. In Proceedings of the International Conference on Image Processing, Santa Barbara, CA, USA, 26–29 October 1997; IEEE Computer Society: Santa Barbara, CA, USA, 1997; Volume 1, pp. 604–607.
24. Zemliachenko, A.N.; Abramov, S.K.; Lukin, V.V.; Vozel, B.; Chehdi, K. Lossy Compression of Noisy Remote Sensing Images with Prediction of Optimal Operation Point Existence and Parameters. *J. Appl. Remote Sens* **2015**, *9*, 095066. [CrossRef]
25. Ponomarenko, N.; Krivenko, S.; Lukin, V.; Egiazarian, K.; Astola, J.T. Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study. *EURASIP J. Adv. Signal Process.* **2010**, *2010*, 976436. [CrossRef]
26. Ponomarenko, N.; Silvestri, F.; Egiazarian, K.; Carli, M.; Astola, J.; Lukin, V. On Between-Corfficient Contrast Masking of DCT Basis Functions. In Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, 13–15 January 2007; 4p.
27. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale Structural Similarity for Image Quality Assessment. In Proceedings of the The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; IEEE: Pacific Grove, CA, USA, 2003; pp. 1398–1402.
28. Zhou, W.; Chen, Z.; Li, W. Dual-Stream Interactive Networks for No-Reference Stereoscopic Image Quality Assessment. *IEEE Trans. Image Process.* **2019**, *28*, 3946–3958. [CrossRef] [PubMed]
29. Cui, Y.; Jiang, G.; Yu, M.; Song, Y. Local Visual and Global Deep Features Based Blind Stitched Panoramic Image Quality Evaluation Using Ensemble Learning. *IEEE Trans. Emerg. Top. Comput. Intell.* **2022**, 1–15. [CrossRef]
30. Zhou, W.; Xu, J.; Jiang, Q.; Chen, Z. No-Reference Quality Assessment for 360-Degree Images by Analysis of Multifrequency Information and Local-Global Naturalness. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 1778–1791. [CrossRef]
31. Lukin, V.V.; Ponomarenko, N.N.; Zelensky, A.A.; Kurekin, A.A.; Lever, K. *Compression and Classification of Noisy Multichannel Remote Sensing Images*; Bruzzone, L., Notarnicola, C., Posa, F., Eds.; SPIE: Cardiff, UK, 2008; p. 71090W.
32. Lukin, V.; Zemliachenko, A.; Abramov, S.; Vozel, B.; Chehdi, K. Automatic Lossy Compression of Noisy Images by Spiht or Jpeg2000 in Optimal Operation Point Neighborhood. In Proceedings of the 2016 6th European Workshop on Visual Information Processing (EUVIP), Marseille, France, 25–27 October 2016; IEEE: Marseille, France, 2016; pp. 1–6.
33. Naumenko, V.; Lukin, V.; Krivenko, S. Analysis of Noisy Image Lossy Compression by BPG. In *Integrated Computer Technologies in Mechanical Engineering—2021*; Nechyporuk, M., Pavlikov, V., Kritskiy, D., Eds.; Lecture Notes in Networks and Systems; Springer International Publishing: Cham, Switzerland, 2022; Volume 367, pp. 911–923, ISBN 978-3-030-94258-8.
34. Kovalenko, B.; Lukin, V.; Naumenko, V.; Krivenko, S. Analysis of Noisy Image Lossy Compression by BPG Using Visual Quality Metrics. In Proceedings of the 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT), Kyiv, Ukraine, 15–17 December 2021; IEEE: Kyiv, Ukraine, 2021; pp. 20–25.
35. BPG Image Format. Available online: https://bellard.org/bpg/ (accessed on 5 May 2022).
36. Yee, D.; Soltaninejad, S.; Hazarika, D.; Mbuyi, G.; Barnwal, R.; Basu, A. Medical Image Compression Based on Region of Interest Using Better Portable Graphics (BPG). In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, USA, 5–8 October 2017; IEEE: Banff, AB, USA, 2017; pp. 216–221.
37. Albalawi, U.; Mohanty, S.P.; Kougianos, E. Energy-Efficient Design of the Secure Better Portable Graphics Compression Architecture for Trusted Image Communication in the IoT. In Proceedings of the 2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI), Pittsburgh, PA, USA, 11–13 July 2016; IEEE: Pittsburgh, PA, USA, 2016; pp. 302–307.
38. Ponomarenko, N.; Lukin, V.; Astola, J.; Egiazarian, K. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013. In *Advanced Concepts for Intelligent Vision Systems*; Battiato, S., Blanc-Talon, J., Gallo, G., Philips, W., Popescu, D., Scheunders, P., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9386, pp. 613–624, ISBN 978-3-319-25902-4.
39. Chiu, M.T.; Xu, X.; Wei, Y.; Huang, Z.; Schwing, A.; Brunner, R.; Khachatrian, H.; Karapetyan, H.; Dozier, I.; Rose, G.; et al. Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. *arXiv* **2020**, arXiv:2001.01306. [CrossRef]
40. Chatterjee, P.; Milanfar, P. Is Denoising Dead? *IEEE Trans. Image Process.* **2010**, *19*, 895–911. [CrossRef] [PubMed]
41. Colom, M.; Buades, A.; Morel, J.-M. Nonparametric Noise Estimation Method for Raw Images. *J. Opt. Soc. Am.* **2014**, *31*, 863. [CrossRef] [PubMed]

42. Ieremeiev, O.; Lukin, V.; Okarma, K.; Egiazarian, K. Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images. *Remote Sens.* **2020**, *12*, 2349. [CrossRef]
43. Ponomarenko, N.; Lukin, V.; Zriakhov, M.; Egiazarian, K.; Astola, J. Lossy Compression of Images with Additive Noise. In *Advanced Concepts for Intelligent Vision Systems*; Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3708, pp. 381–386, ISBN 978-3-540-29032-2.
44. Ponomarenko, N.; Zriakhov, M.; Lukin, V.V.; Astola, J.T.; Egiazarian, K.O. Estimation of Accessible Quality in Noisy Image Compression. In Proceedings of the 2006 14th European Signal Processing Conference, Florence, Italy, 4–8 September 2006; pp. 1–4.
45. Pogrebnyak, O.; Lukin, V.V. Wiener Discrete Cosine Transform-Based Image Filtering. *J. Electron. Imaging* **2012**, *21*, 043020. [CrossRef]
46. Cameron, C.A.; Windmeijer, F.A.G. An R-Squared Measure of Goodness of Fit for Some Common Nonlinear Regression Models. *J. Econom.* **1997**, *77*, 329–342. [CrossRef]
47. Rousseeuw, P.J. Least Median of Squares Regression. *J. Am. Stat. Assoc.* **1984**, *79*, 871–880. [CrossRef]
48. Lukac, R.; Plataniotis, K.N. (Eds.) *Color Image Processing: Methods and Applications*; Image Processing Series; CRC/Taylor & Francis: Boca Raton, FL, USA, 2007; ISBN 978-0-8493-9774-5.
49. Ponomarenko, N.; Lukin, V.; Egiazarian, K.; Astola, J. DCT Based High Quality Image Compression. In *Image Analysis*; Kalviainen, H., Parkkinen, J., Kaarna, A., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3540, pp. 1177–1185, ISBN 978-3-540-26320-3.
50. Pyatykh, S.; Hesser, J.; Zheng, L. Image Noise Level Estimation by Principal Component Analysis. *IEEE Trans. Image Process.* **2013**, *22*, 687–699. [CrossRef]
51. Zoran, D.; Weiss, Y. Scale Invariance and Noise in Natural Images. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 27 September–4 October 2009; IEEE: Kyoto, Japan, 2009; pp. 2209–2216.
52. Rodrigues, I.; Sanches, J.; Bioucas-Dias, J. Denoising of Medical Images Corrupted by Poisson Noise. In Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008; IEEE: San Diego, CA, USA, 2008; pp. 1756–1759.
53. Zhang, B.; Fadili, J.M.; Starck, J.L. Wavelets, Ridgelets, and Curvelets for Poisson Noise Removal. *IEEE Trans. Image Process.* **2008**, *17*, 1093–1108. [CrossRef] [PubMed]