# Measuring objective image and video quality: On the relationship between SSIM and PSNR for DCT-based compressed images

Maria G. Martini, *Senior Member, IEEE*

***Abstract*—Measuring accurately image and video quality is a critical step in any image and video processing and compression method and streaming / broadcasting system. In particular, simple and tractable objective metrics are required for quality-driven system optimization. The aim of this paper is to show how the structural similarity metric (SSIM) for image quality assessment can be seen in many cases, such as Discrete Cosine Transform (DCT)-based compressed images and video, as a content-aware version of the peak signal-to-noise ratio (PSNR) and it can be accurately estimated based on it. In fact, under some assumptions described in the paper, the first can be derived directly from the latter based on a single content-dependent parameter, i.e. the variance of the image / video frame. Tests on example images compressed via JPEG at different quality levels further validate the assumptions and show how the proposed derivation can be utilized in replacement of the original expression of SSIM for compressed images/video frames at quality levels of interests in real applications (e.g., video streaming). Finally, as an example application of the derivation, we derive an expression for measuring the image/video quality following image/video transcoding quality based on SSIM.**

***Index Terms*—Quality assessment, objective quality metrics, PSNR, SSIM, JPEG, image and video compression**

## I. Introduction

Image/video compression is a key step in any video streaming system, enabling the reduction of the required transmission data rate. Lossy image/video compression results, however, in a reduction of image/video quality. The quality of impaired images can be assessed via subjective tests, with a pool of subjects visualising them and providing a score in a specific scale (e.g., 5-point Likert scale); the mean of the scores from the pool of subjects (mean opinion score, MOS) is typically considered as the quality value associated to a specific visual stimulus. This procedure, however, is expensive in terms of time of subject viewers, logistic resources (use of specialised labs) and time to collect and analyse the scores. An alternative method is to use objective quality metrics, calculated via mathematical expressions or algorithms or learning-based models. Such metrics can be calculated comparing the impaired and original image/video (full reference), or based only on the impaired image/video (no-reference) or based on the impaired image/video and using only limited information from the original image/video (reduced-reference).

Figure 1 depicts via a block diagram the process of lossy compression and the relevant objective, full reference quality assessment (bottom) based on original and reconstructed image or video.

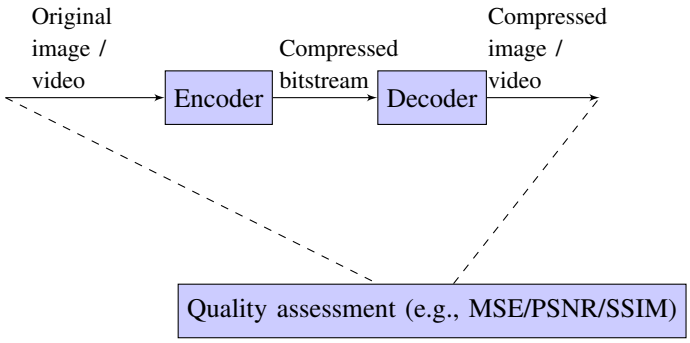



Fig. 1: Quality assessment as comparison between original and compressed video.

Focusing on a single image or video frame, the Mean Square Error (MSE) between the original (X) and the compressed version (Y) is calculated as follows:

$$MSE = \frac{1}{MN} \sum_{j=1}^{M \times N} (x_j - y_j)^2 = \mathcal{E}(e^2) \tag{1}$$

where $M$ and $N$ are the number of pixels in horizontal and vertical direction, $e$ represents the error and $\mathcal{E}$ is used to identify the mean.

The Peak Signal to Noise Ratio (PSNR) is derived from MSE as follows:

$$PSNR = 10 \log_{10} \frac{(2^b - 1)^2}{MSE} \tag{2}$$

where $b$ is the bit depth (number of bits per pixel).

While MSE and PSNR are easy to calculate, they are not always good indicators of the actual quality as perceived by the users. For this reason, other quality metrics have been developed. In particular, Structural Similarity Index (SSIM) [1] has become very popular, since it has been shown to correlate well with the quality as perceived by humans for different types of distortions.

The SSIM metric is defined as below:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

where $\mu_x$ and $\mu_y$ are the mean of the original and the impaired grayscale image, respectively, while $\sigma_x^2$ and $\sigma_y^2$

Maria Martini is with Kingston University London, UK.
Manuscript received August 27, 2024; accepted November 7, 2024.

are the variance of the original and the impaired grayscale image, respectively; $\sigma_{xy}$ represents the covariance between original and impaired images (grayscale) and $C_1$ and $C_2$ are constant values to make sure the metric is a real number. A similar index can be calculated on the color components of an image. SSIM is a full reference metric and it can be decomposed in terms addressing structure, contrast, and luminance comparison.

PSNR and SSIM are still the most widely used metrics for image and video quality assessment [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15], being also used for rate control [16] [17] and rate-distortion optimisation [18] [19] [20] [21] [22], as well as super-resolution strategies [23] [24] describing quality levels of content in image/video datasets [25] [26] [27] [28] [28] [29] [30].

For some of these purposes (e.g., in-loop rate-distortion optimisation) a mathematically tractable metric is required, since subjective quality metrics cannot be used directly and the same applies to objective quality metrics based on machine learning (e.g., Video Multimethod Assessment Fusion (VMAF) [31], Learned Perceptual Image Patch Similarity (LPIPS) [32]). While SSIM better reflects human perception and may appear mathematically tractable, it is more complex than MSE and PSNR and requires the calculation of the covariance of original and impaired image, i.e., joint knowledge and use of all the pixel values of original and impaired image/video, besides mean and variance of both original and impaired image/video.

In order to simplify its calculation, this paper discusses the derivation of SSIM directly from PSNR or MSE in the case of Discrete Cosine Transform (DCT)-based image and video compression. The latter case allows simplifications since it makes possible the assumption that the mean of the luminance/chrominance values does not vary with the compression ratio and the same assumption can be made - to some extent - for the variance.

The contributions of this paper include:

- A detailed analytical derivation and discussion of the relationship between SSIM and MSE/PSNR for DCT-based compressed images and video.
- Following the observation that mean and variance of an image or video frame are not evidently affected by compression (at the bitrates of interest), two accurate and mathematically tractable approximations are presented for the calculation of SSIM as a function of MSE/PSNR, for different use cases. Under the considered assumptions, the two quality metrics are related via a simple content dependent parameter (the variance of either of the compressed or original image), hence SSIM can be seen as a content-aware PSNR, not needing the calculation of the covariance between the two images as in the original expression.
- In particular, in the case the variance of the original image is not available (e.g., image/video delivered over a network), the paper presents an accurate way to calculate SSIM based on only MSE or PSNR and the variance of the compressed image as only content-dependent factor. An advantage of this method is that, assuming PSNR information but not SSIM is provided, for instance as metadata in a received bitstream, SSIM can be calculated easily with no need for other reference to the original video, with a mathematically tractable model that can be used in the formulation of system optimization.
- In the case of (block-based) rate-distortion optimization for image and video compression, the paper suggests the use of the SSIM approximation with the variance of the original image, available in this case. With the presented accurate SSIM approximation, only the source variance has to be computed once per block, while the co-variance would require computation for the reconstruction of a block for each possible mode decision.
- The comparison with another approach in the literature is also presented. The approach, considered as benchmark, aimed to derive a more realistic MSE-based metric for compressed images, also in link with SSIM, with different assumptions. It is shown that the model presented here is more accurate to estimate SSIM in the case of DCT compressed images/video frames, as detailed in the results section.
- Assumptions done in previous SSIM-based rate-distortion optimisation strategies are discussed in light of the observations provided in this paper.
- Finally, as an example application of the derivation, a new simple expression for image/video transcoding quality based on SSIM is provided.

This paper is a major extension of [33].

The remainder of this paper is organized as follows. After a summary of related work in Section II, an analysis of the relationship between SSIM and PSNR is presented in Section III, while Section IV provides comparative results for example compressed images. As an example application of the derivation, Section V proposes a method to assess the perceptual quality of transcoded video based on MSE. Section VI presents concluding remarks along with a brief discussion of further potential applications.

## II. Related work

A review of the state of the art in the field including the main statistical performance evaluation methods is provided in [34].

There have been a number of attempts to compare PSNR and SSIM and both metrics with the results of subjective tests summarised via the Mean Opinion Score (MOS). Raw approximations, not taking image / video content into account, resulted in tables (see e.g. [35]), used for the design and optimization of multimedia systems. Focusing on quality reduction caused by packet losses, in [36] the authors mentioned and used a quite simple relationship for block-based SSIM as a function of MSE, information on original and impaired image, and constant terms. The relationship between SSIM and PSNR has been further analysed in [37] [38] where analytical expressions and approximations are provided for different use cases. According to the approximations considered, the calculation of SSIM from PSNR still requires some joint processing of original and compressed image, beyond

what required for PSNR. In particular, in [37] the proposed relationship involves the evaluation of the covariance between original and impaired image and a linear approximation is proposed for SSIM values between $0.2$ and $0.8$ (too low for use cases of practical interest). Such relationship is explored with examples in [38]. This work also highlights how the luminance comparison component in SSIM has a marginal impact on the final value for the examples considered, while the structure comparison term has a higher impact than the contrast comparison one. In [39] the authors, while discussing the validity of SSIM as quality metric, also showed that the index is directly related to the mean squared error, compared the two metrics statistically and derived a pair of functions that algebraically connect the two via means and mean square values of original and impaired images. Information on both original and impaired image is requested by their formulation in addition to MSE.

The most appropriate values for the different constants used in the SSIM equations are analysed empirically in [40], together with the window size in the calculation of MSSIM, in order to increase the correlation with MOS.

In [41] the associations between MSE and SSIM as cost functions in linear decomposition are investigated. It is observed that in this case the selected bases from a basis set for a target vector are the same in the linear decomposition schemes with different cost functions MSE and SSIM. In addition, for a target vector, the ratio of the corresponding linear coefficients of the selected bases in the MSE-based linear decomposition scheme and the SSIM-based scheme is a constant and corresponds to the value of the Pearson's correlation coefficient between the target vector and its estimated vector.

The authors of the work in [42], after recognising that the SSIM expression can be simplified to a correlation coefficient in the case $\mu_x = \mu_y$, propose an alternative quality metric based on similar statistics. Other works (such as [43]) aim at estimating the SSIM metric in absence of a reference from bitstream and/or reconstructed video and compare the estimation with PSNR estimation, but not establishing a relationship between the two. Two of the authors of the SSIM metric discuss in [44] the properties of MSE and SSIM, but they do not focus explicitly on their relationship. Bounds on the SSIM index are provided for compressed images in an interesting work [45] as a function of quantization rate for uniform, Gaussian, and Laplacian sources. The work proposes to use the bounds for rate allocation problems in practical image and video coding applications. No relationship with PSNR or MSE is established. In [46] the authors proposed a reduced-reference (RR) metric based on SSIM, highlighting an interesting linear relationship between the full reference SSIM measure and their RR estimate when the image distortion type is fixed. The work does not focus however on the relationship with PSNR or MSE. In [47] the authors developed a perceptually relevant MSE-based image quality metric. In doing so, they assume an additive error model and independence between signal and error. This metric is also adopted in [18] and used as a benchmark in the results section of this paper. It is shown below that the estimation provided is not as accurate as those proposed and discussed here for DCT compressed images. The lack of accuracy of the estimation was also reported in [22]. The mathematical properties of the structural similarity index have been studied in [48] and finally a very recent work [49] provides an alternative model of SSIM computation (utilizing subband decomposition and identical distance measures in each subband) and discusses relationship with PSNR, highlighting that the two are linked via a joint statistic of both signals (covariance).

In this paper this term is simplified, in the scenario and under the assumptions considered. The assumptions also allow an accurate estimation of SSIM based on information on a single version of the image (either original or compressed) besides MSE. This is useful, e.g., for real time rate-distortion optimisation for image/video compression or reduced reference image quality assessment when metadata on MSE or PSNR is provided.

## III. Analysis of the relationship between SSIM and PSNR

In order to identify a simple relationship between SSIM and PSNR, it is assumed that for distortion due to DCT-based compression $\mu_e = 0, \mu_x = \mu_y$ (equivalent) and $\sigma_x^2 = \sigma_y^2$. Similar assumptions were considered in other works for other purposes [50] [51].

To show the validity of this assumption with an example, Figures 2a and 3a report mean and variance for the Baboon sample image compressed according to the Joint Photographic Experts Group (JPEG) standard [52] (still the most popular image compression standard) at different compression ratios (associated to the quality factors reported in the horizontal axis). Figures 2b and 3b report mean and variance for the Peppers sample image.

We can observe that indeed the mean of the luminance is approximately constant when the image is compressed at widely different compression ratios / quality factors (the mean - of the possible luminance values between 0 and 255 - ranges from 145.83 to 145.94 with an extremely limited variation, within 0.075%). According to Figure 3a the variation of the global variance with the quality factor is limited between 2200 and 2400 (variation within 9%). A similar behavior can be observed for the Pepper sample image (Figures 2b and 3b).

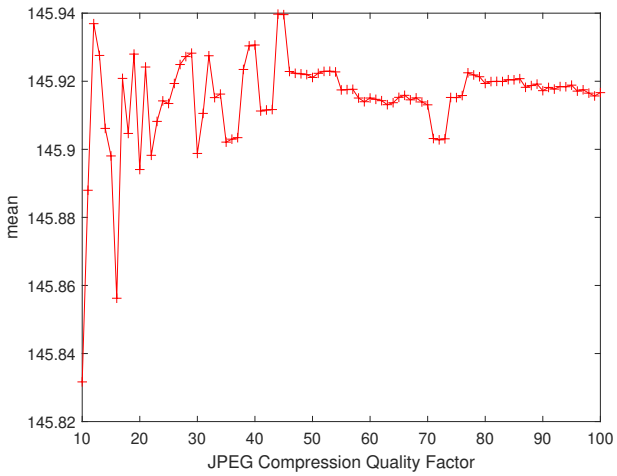
To show the validity of the assumptions also at local value, block-based variance for the Baboon image is reported in Fig. 4. Each curve in the figure represents the variance of a different block (the first five blocks are considered to make the figure readable) and shows its variation with the compression level. We can see how variation is very limited also at local level.

Using only the first assumption on the mean we have from (3):

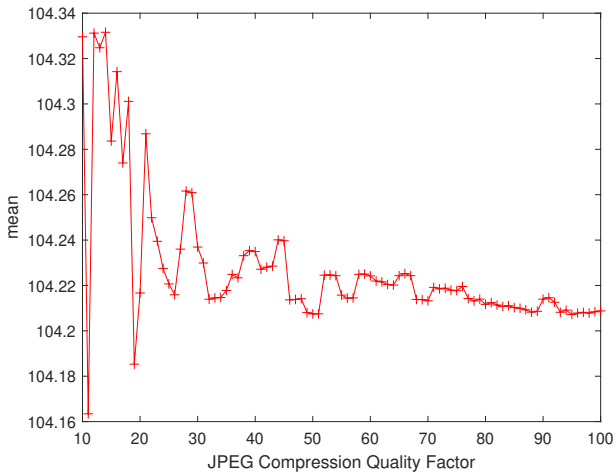
$$\mathrm{SSIM}(x,y) = \frac{(2\mu_y^2 + C_1)(2\sigma_{xy} + C_2)}{(2\mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{4}$$

hence:

$$\mathrm{SSIM}(x,y) = \frac{(2\sigma_{xy} + C_2)}{(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{5}$$

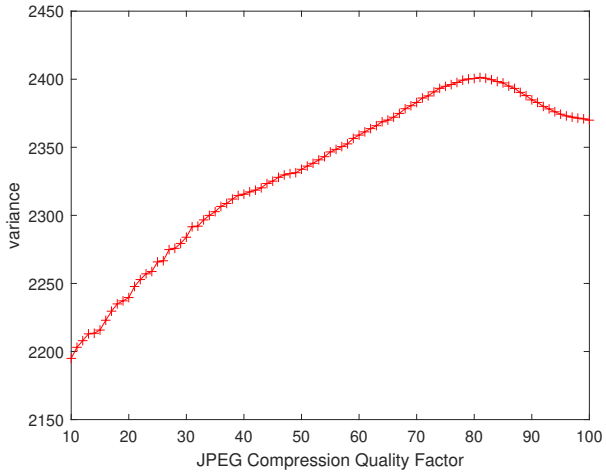



(a) Baboon image

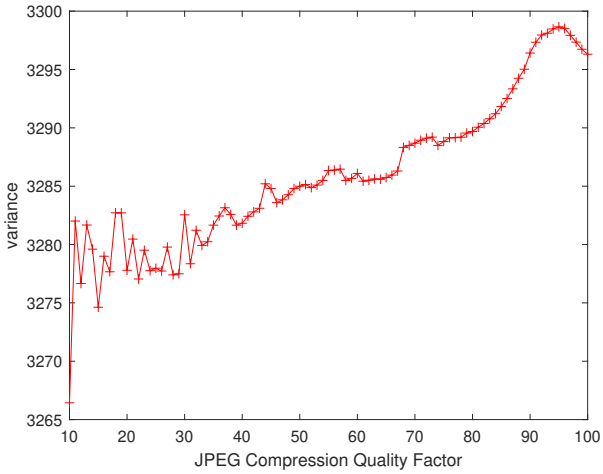



(b) Peppers image

Fig. 2: Mean for two example images at different JPEG compression quality factors




(a) Baboon image




(b) Peppers image

Fig. 3: Variance for two example images at different JPEG compression quality factors.

We also have:

$$e = x - y \tag{6}$$

$$\sigma_e^2 = \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy} \tag{7}$$

hence:

$$2\sigma_{xy} = \sigma_x^2 + \sigma_y^2 - \sigma_e^2. \tag{8}$$

Using (8) in (4)

$$\text{SSIM}(x,y) = \frac{(\sigma_x^2 + \sigma_y^2 - \sigma_e^2 + C_2)}{(\sigma_x^2 + \sigma_y^2 + C_2)} = 1 - \frac{\sigma_e^2}{(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{9}$$

With the previous assumption $\mu_e = 0$ we have

$$\sigma_e^2 = \mathcal{E}(e^2) = MSE \tag{10}$$

hence:

$$\text{SSIM}(x,y) = 1 - \frac{MSE}{(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{11}$$

Using also the assumption

$$\sigma_x^2 = \sigma_y^2 : \tag{12}$$

$$\text{SSIM}(x,y) = 1 - \frac{MSE}{(2\sigma_y^2 + C_2)}. \tag{13}$$

The reader can note that in (11) and (13) there is no dependency on the covariance, that requires full information on original and compressed image for evaluation. Also, $C_2$ is a constant (the same one as in the SSIM formulation). The result is consistent with previous observations [19] [53] of linear relationship between coding rate $R$ and $MSE$ and coding rate $R$ with $1 - SSIM$. Some authors intuitively

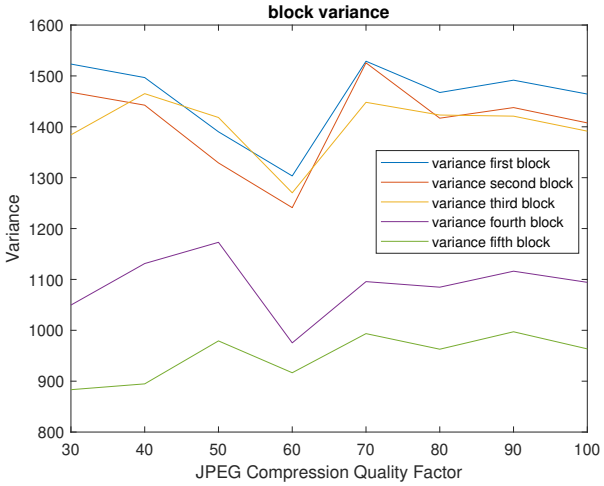



Fig. 4: Block-based variance for the example Baboon image at different JPEG compression quality factors (first five blocks).

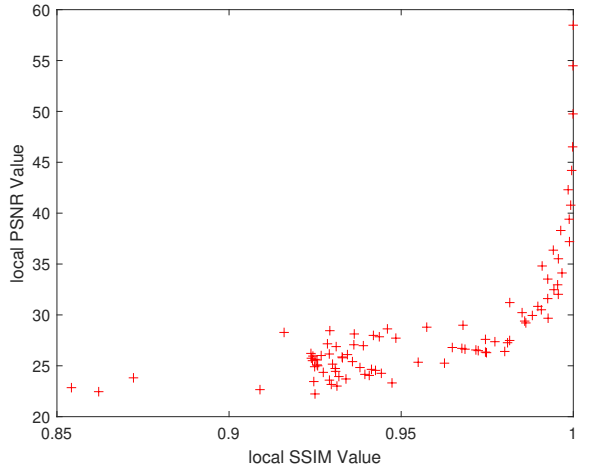



Fig. 5: PSNR vs. SSIM for the first block of the Baboon example image at different JPEG compression levels (Quality Factor from 10 to 100).

used $1 - SSIM(x, y)$ as distortion measure in rate-distortion optimization, since $1 - SSIM(x, y)$ represents a distortion value while SSIM represents a quality value [54] [53] [55]. In particular, the authors of [53] considered a hybrid optimization method aiming at jointly minimising MSE and 1-SSIM. The two are however tightly related and the relationship between the two is clear from equation (13) that can be written as:

$$1 - \text{SSIM}(x, y) = \frac{MSE}{(2\sigma_y^2 + C_2)} \tag{14}$$

or, using equation (12),

$$1 - \text{SSIM}(x, y) = \frac{MSE}{(2\sigma_x^2 + C_2)}. \tag{15}$$

Considering PSNR:

$$PSNR = 10log_{10}\frac{(2^b - 1)^2}{MSE} \tag{16}$$

we can write:

$$MSE = \frac{(2^b - 1)^2}{10^{PSNR/10}} \tag{17}$$

and using (17) in (13) we have:

$$\text{SSIM}(x, y) = 1 - \frac{\frac{(2^b-1)^2}{10^{PSNR/10}}}{(2\sigma_y^2 + C_2)}. \tag{18}$$

It can be noted that this is also a way to calculate SSIM with reduced computational complexity: rather than needing to calculate the mean and variance of both original and impaired image, as well as the covariance, only the variance of the impaired image needs to be calculated, in addition to PSNR (18) or MSE (13). Since PSNR is content-independent, the variance of the impaired image is the only element taking into account the content of the image under assessment.

The global structural similarity index is typically calculated as the mean of the SSIM of sub-windows composing the image:

$$\text{MSSIM}(X, Y) = \frac{1}{M}\sum_j \text{SSIM}(x_j, y_j) \tag{19}$$

where $X$ and $Y$ are the reference and the distorted images, respectively, $x_j$ and $y_j$ are the image contents at the $j$-th local window, and $M$ is the number of samples in the quality map [1].

Hence,

$$\text{MSSIM}(X, Y) = \frac{1}{M}\sum_j [1 - \frac{\frac{(2^b-1)^2}{10^{PSNR_j/10}}}{(2\sigma_{yj}^2 + C_2)}]. \tag{20}$$

## IV. Numerical results

This section reports example results with images frequently used in the image processing community, to highlight the validity of the approximations considered while enabling reproduction of the results. The six images considered are grayscale images widely used in the research community, with resolution $512 \times 512$, covering different types of contents and complexities (Baboon, Lena, Barbara, Peppers, Goldhill, Boat).

An example of local (block-based) results are provided in Figure 5, showing the scatter plot for local PSNR value vs. SSIM value for the same (first) block of the Baboon image. A square block of $8 \times 8$ pixels is considered for these results.

In the following, results are provided as the mean of block-based results across the image (MSSIM) as in eq. (20). For the comparison below, the Matlab *ssim index* implementation has been considered, with $16 \times 16$ window. MSE is calculated on the same size window, according to eq. (18), and the results are then averaged according to (19). Figures 6, 7, 8, 9, 10, 11 report in the top row the comparison of the MSSIM value

(blue curve), calculated with the original expression (*ssim index* code associated with [56]), with:

- $MSSIM_{MSE}$ - MSSIM calculated based on mean local MSE as in eq. (11) and eq. (19) considering only the assumption $\mu_x = \mu_y$ (magenta curve);
- $MSSIM_{MSE,C}$ - MSSIM calculated based on mean local MSE as in eq. (13) and eq. (19) considering both assumptions $\mu_x = \mu_y$ and $\sigma_x^2 = \sigma_y^2$ and considering $\sigma_y^2$ (variance of the compressed image) for the calculation (red curve).
- $MSSIM_{MSE,O}$ - MSSIM calculated based on mean local MSE as in eq. (13) and eq. (19) considering both assumptions $\mu_x = \mu_y$ and $\sigma_x^2 = \sigma_y^2$ but considering $\sigma_x^2$ (original image) for the calculation (green curve).

To facilitate understanding, the subscripts O for original and C for compressed are used in the names above and in the figures legend, to denote that the variance of the original image or the variance of the compressed image were used in the computation, respectively. It can be observed that, for the typical quality range of interest, the proposed function of MSE approximates the MSSIM score with high accuracy. In particular $MSSIM_{MSE}$ (magenta curve) is almost indistinguishable from the MSSIM calculated with the original expression (blue curve) in all cases considered (in some cases the blue and magenta curves are completely overlapped and only one is visible). In this case, the only assumption made is that the means of the original and compressed images are equivalent ($\mu_x = \mu_y$) while no assumption is made on the variance. $MSSIM_{MSE,C}$ (red curve), obtained with the variance of the compressed image, provides an excellent approximation of MSSIM calculated with the original expression (blue curve) in all cases and tends to slightly underestimate it, as expected (the variance in general is lower for compressed images and gets lower as the quality factor / bitrate decreases, as can be observed in the bottom part of Figures 6, 7, 8, 9, 10, 11 Interestingly, this happens at PSNR values (25-28) below the PSNR values of typical interest, where the approximation holds tightly. $MSSIM_{MSE,O}$ (green curve) also provides an excellent approximation of MSSIM calculated with the original expression (blue curve) in all cases and tends to slightly overestimate it, as expected. In fact, the actual SSIM expression uses both $\sigma_x^2$ and $\sigma_y^2$ while here both are replaced with $\sigma_x^2$, in general higher than $\sigma_y^2$. Also in this case, for more complex contents (e.g., Baboon), it can be observed that for low JPEG quality factor values (and low PSNR - see second row of the figures) the difference becomes more evident. The bottom row of Figures 6, 7, 8, 9, 10, 11 reports the mean local variance of the compressed image as a function of the JPEG quality factor, showing the validity of the assumption for most contents and also how the mean local variance is related to the complexity of the content. This observation is used in [57] to estimate image and video content complexity.

Figures 12 and 13 report the comparison in the case of JPEG compression with the MSE-based metric proposed in [47], that assumes an additive distortion model, where the error is assumed uncorrelated with the signal:

$$\frac{1}{SSIM_{fromMSE}} = 1 + \frac{MSE}{2\sigma_x^2 + C_2}. \tag{21}$$

The reader can note that this expression uses the variance of the original image, hence can be applied at the coding/transmission side, but not as reduced-reference metric. Since this model relies on the variance of the original image, it can be directly compared with (15). Note that the error due to compression is content-related, hence the assumption considered in [47] is hardly valid in our case of DCT compressed images. Indeed, the metric proposed in [47] (blue curve with +) is a worse approximation of the actual MSSIM values (blue curve with o) with respect to those proposed here in the case of availability of the variance of the original image (green and magenta curves). Referring to the model in [47][18], also [22] reports that its modeling accuracy for High Efficiency Video Coding (HEVC) is less than satisfactory.

To further analyse the comparison between the two, it can be observed that the term

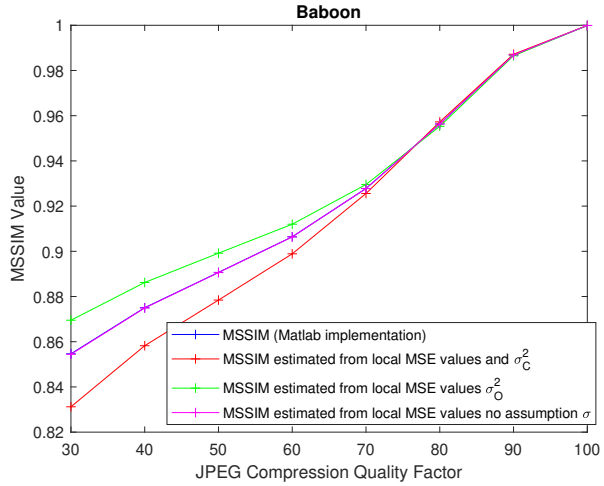
$$\frac{MSE}{2\sigma_x^2 + C_2} \tag{22}$$

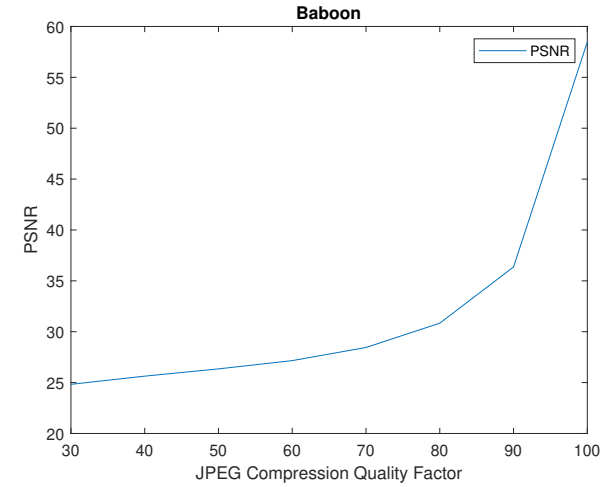
is considered / approximated as equivalent in (21) to $(1/SSIM) - 1$ and in (15) to $1 - SSIM$. Figure 15 shows how these two relate, for the range of SSIM values of interest.

As briefly anticipated above, some works focusing on SSIM-based rate-distortion optimization used either (1-SSIM) [19] [16] or (1/SSIM) [18] as distortion metric. The discussion presented above in this paper illustrates and clarifies how these two are related between them and to MSE.

Figure 14 reports the results in case of video coding. H.264 coding (with "slow" preset) was used, at different Constant Rate Factor (CRF) values (this is reflecting current systems: when CRF is selected the bitrate is adjusted based on the complexity of the content and low CRF values result in higher quality). The left panel shows the results for an Intra (I) frame (the first one), while the right one shows the results for a predicted frame (the second one). The Foreman video sequence (in CIF format) was considered here to facilitate comparisons. We can observe that the results in the case of video show a globally similar behaviour to those obtained for images, also for predicted frames, where encoding is based on the calculation of motion vectors and motion compensated differences. We remark that, when the variance of the original image is available, the metric proposed here with $\sigma_O$ definitely outperforms the benchmark. When the variance of the original image is not available, the benchmark cannot be used and the proposed approximation with $\sigma_C$ provides an accurate approximation also in this case. We note that in this case, although excellent, the approximation with no assumption on the variance is not completely overlapped with the MSSIM curve, which means that the assumption of equal mean - only assumption done in this case - is not as precise as in the previous case, although still accurate.

Future work will include methods for variance estimation and fast evaluation, to further improve the approximation accuracy and computational speed.



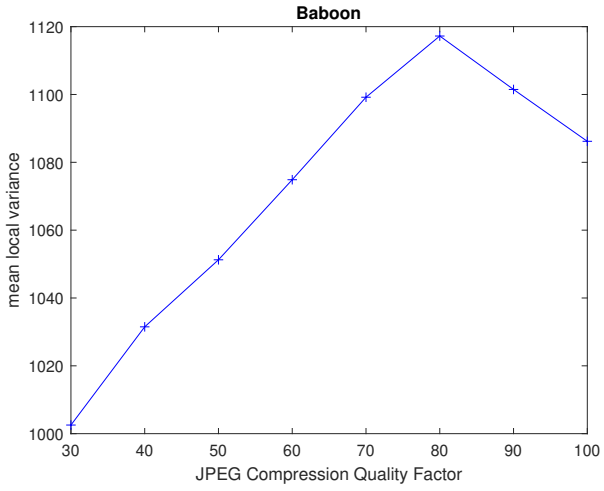
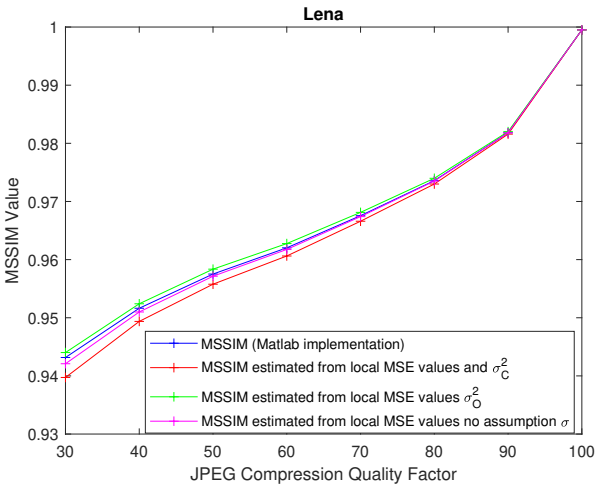
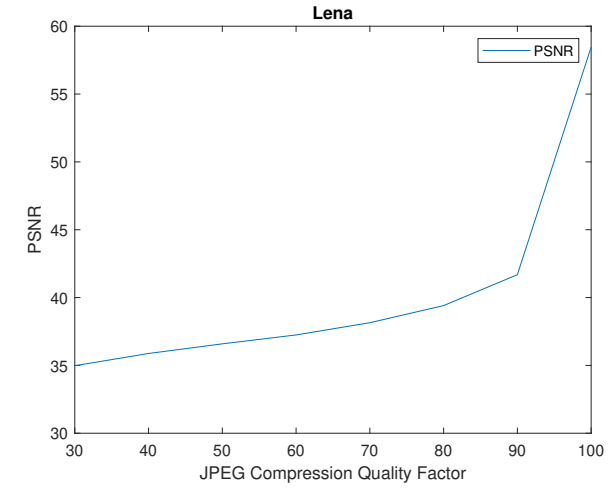




Fig. 6: Mean Structural Similarity Index (MSSIM) (top row), PSNR (mid-row), and mean local variance (bottom row) for Baboon image compressed with JPEG at different quality levels.



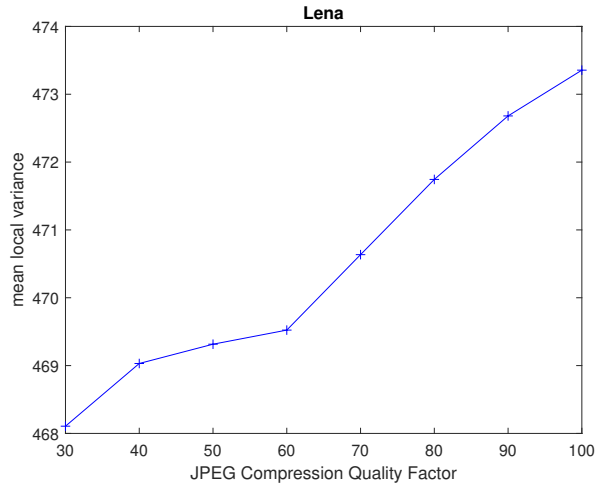
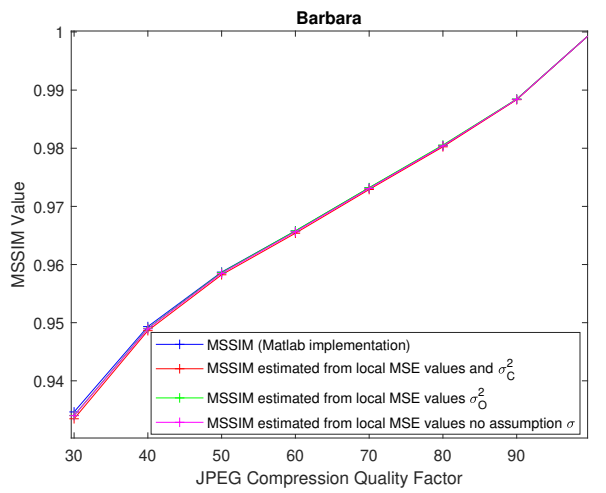
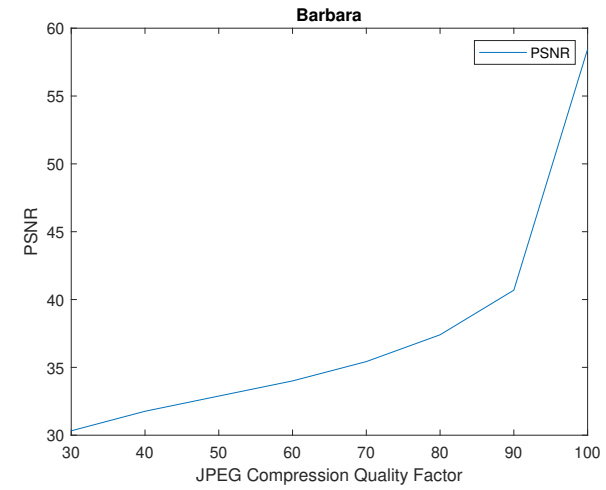




Fig. 7: MSSIM (top row), PSNR (mid-row), and mean local variance (bottom row) for Lena (b) image compressed with JPEG at different quality levels.



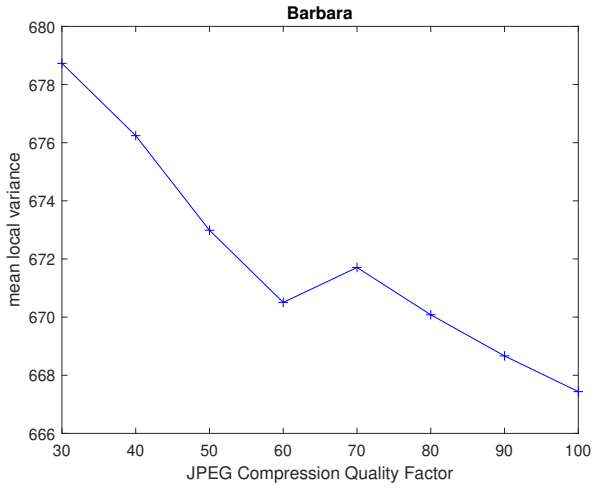
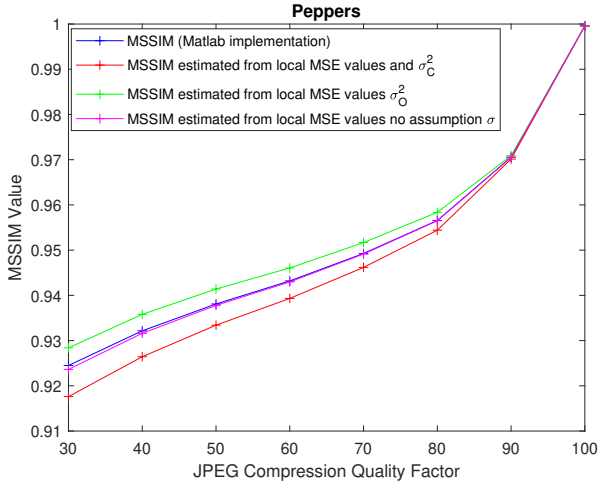
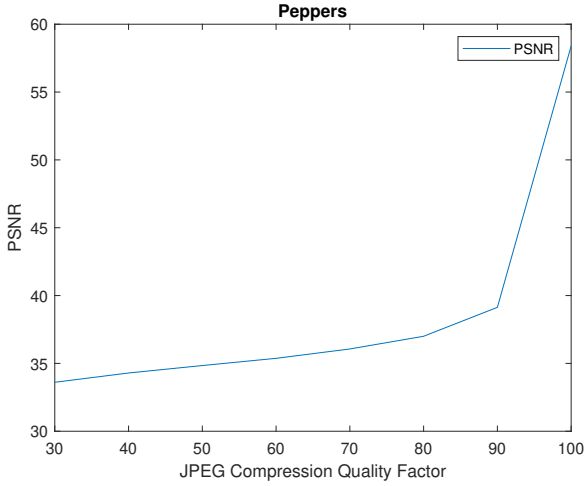




Fig. 8: MSSIM (top row), PSNR (mid-row), and mean local variance (bottom row) for Barbara image compressed with JPEG at different quality levels.



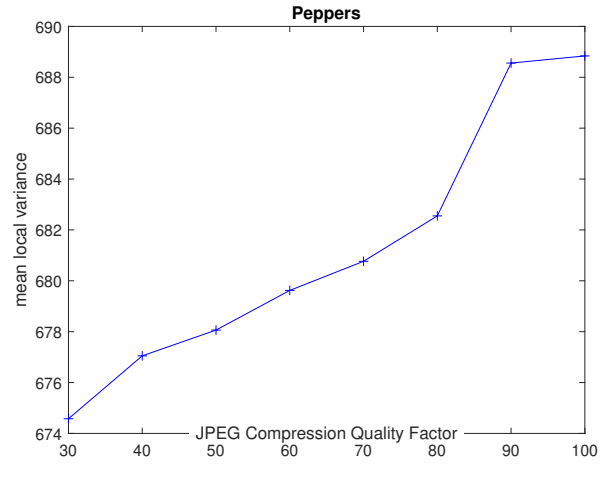
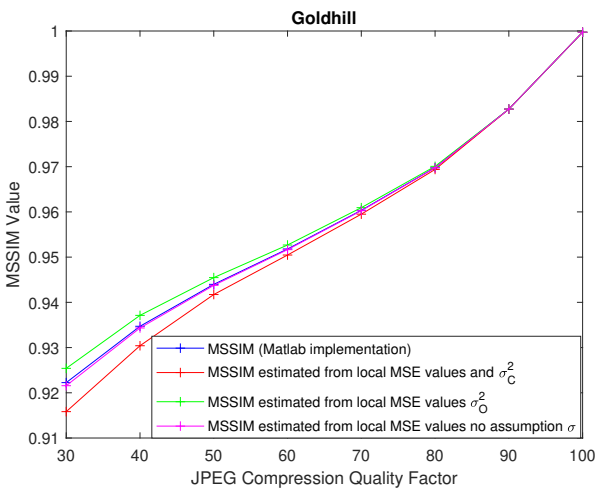
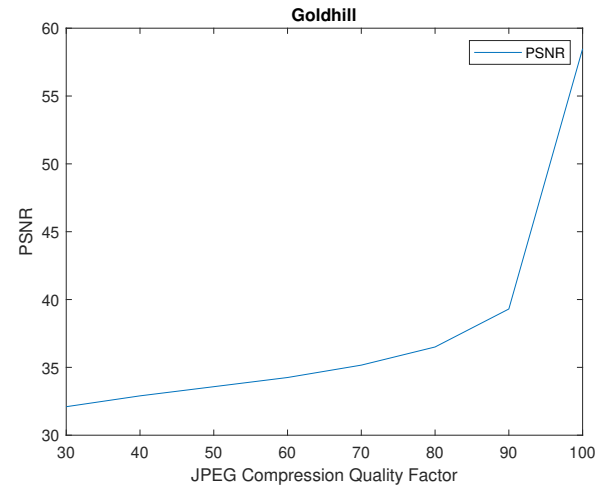




Fig. 9: MSSIM (top row), PSNR (mid-row), and mean local variance (bottom row) for Peppers image compressed with JPEG at different quality levels.



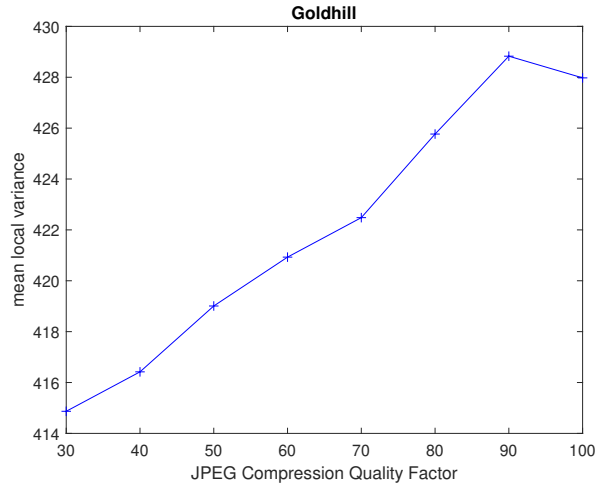
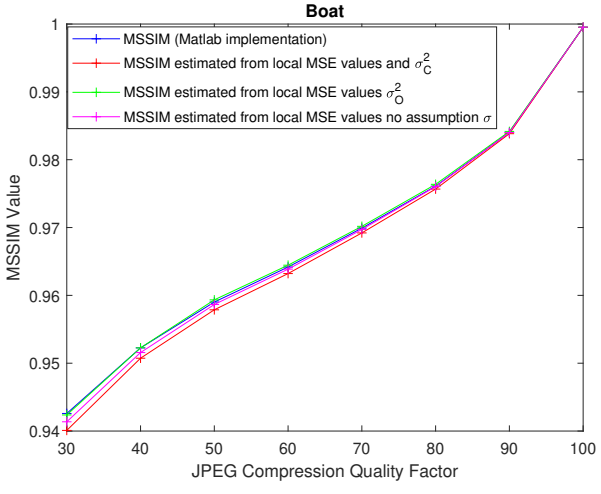
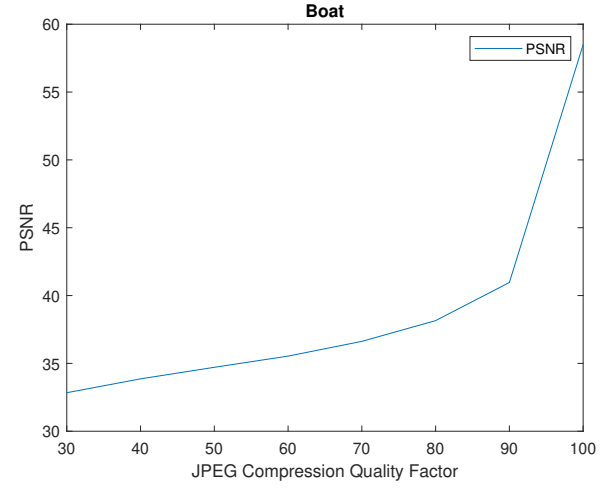




Fig. 10: MSSIM (top row), PSNR (mid-row), and mean local variance (bottom row) for Goldhill image compressed with JPEG at different quality levels.



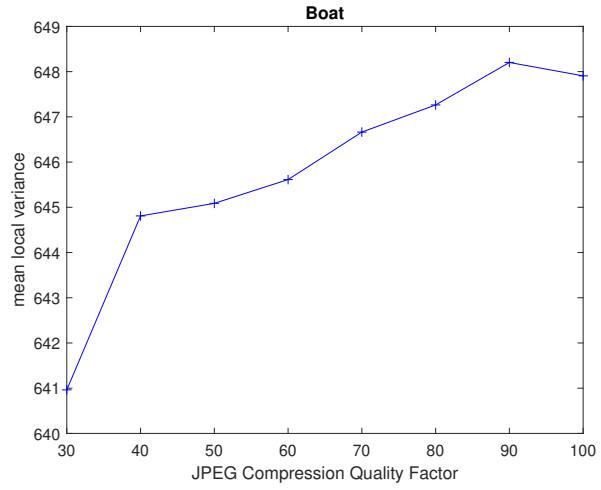
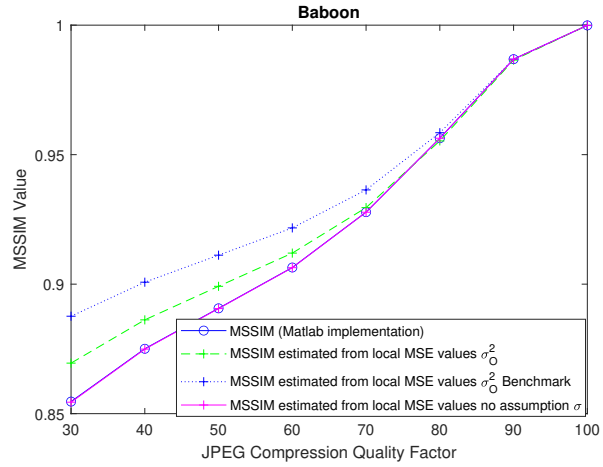
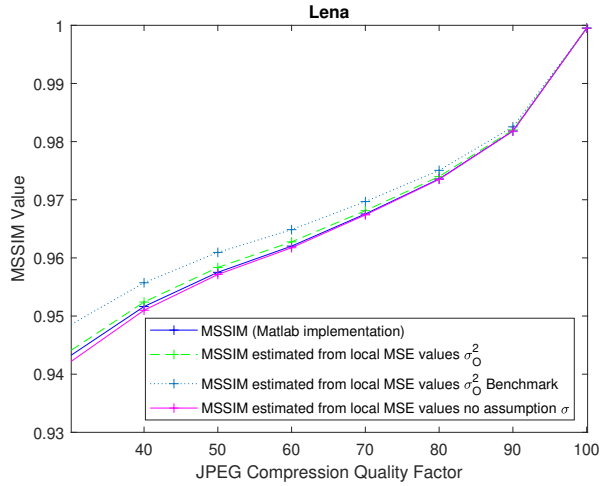




Fig. 11: MSSIM (top row), PSNR (mid-row), and mean local variance (bottom row) for Boat image compressed with JPEG at different quality levels.



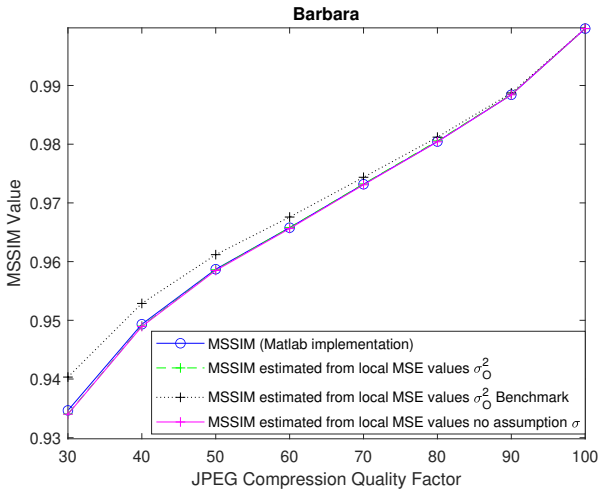
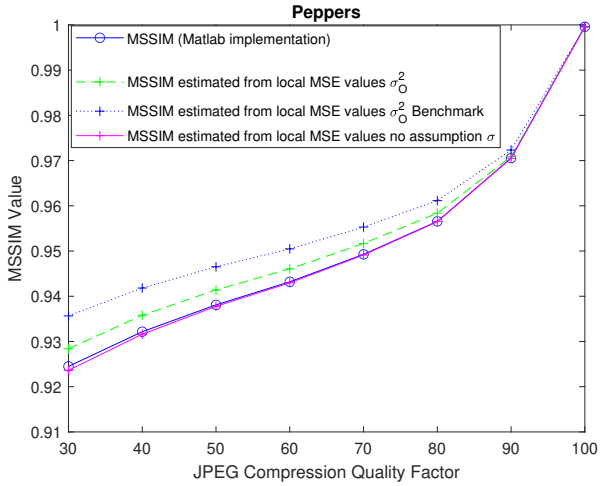
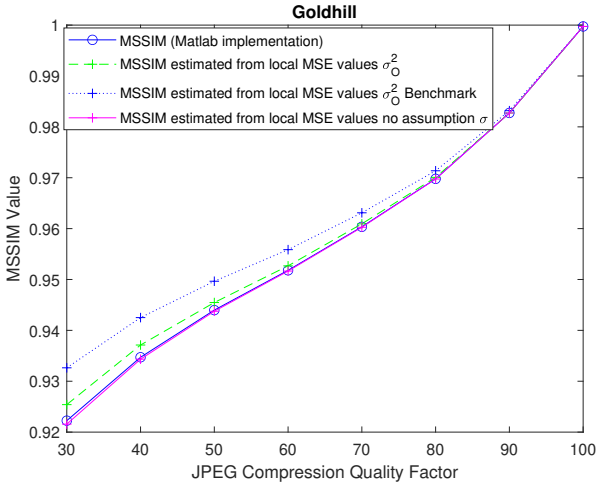




Fig. 12: Comparison with metric in [47].



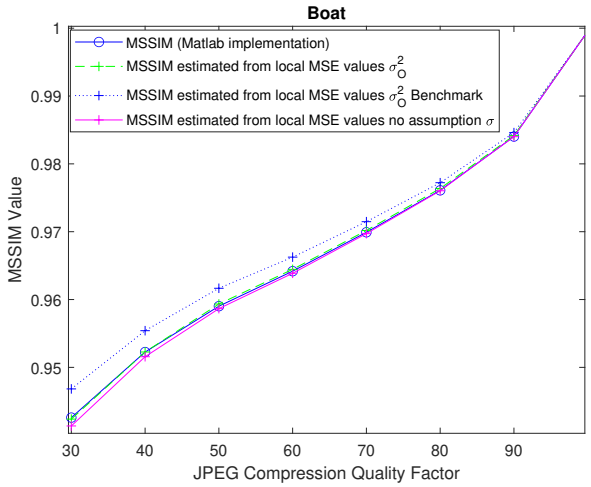
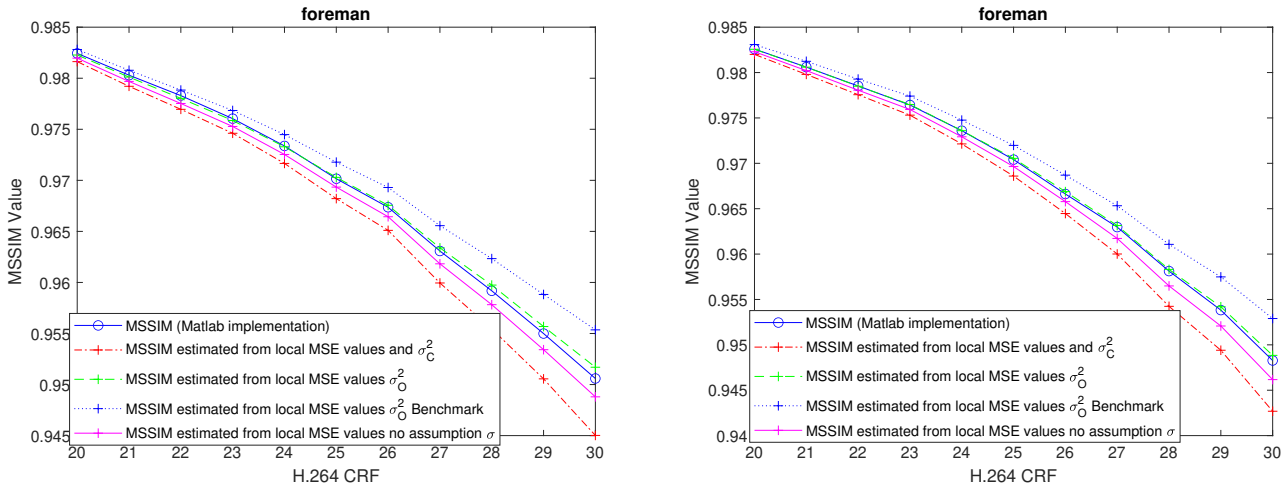




Fig. 13: Comparison with metric in [47].




(a) Intra frame

(b) Predicted frame

Fig. 14: Performance results for H.264 video compression at different CRF values, including comparison with the metric in [47].

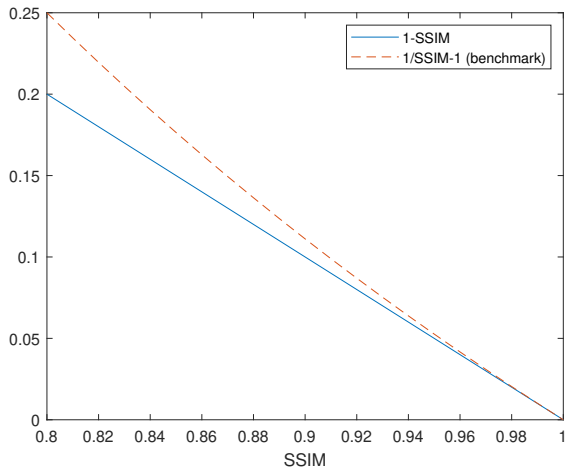



Fig. 15: Comparison between the terms $(1/SSIM) - 1$ and $1 - SSIM$ in (21) and (15), respectively.

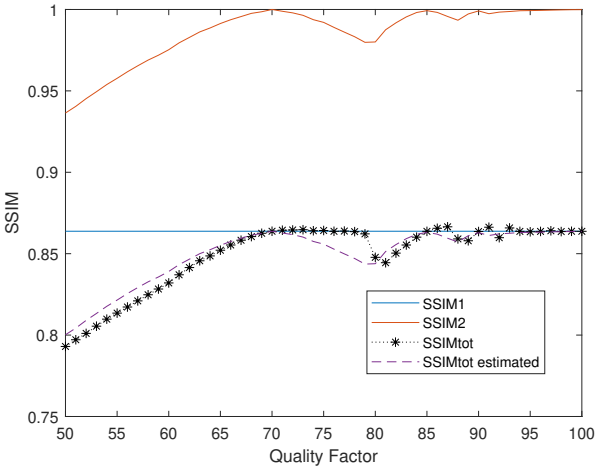



Fig. 17: SSIM for Baboon image, encoded first at QF 70, then transcoded with QF from 50 to 100, default Gaussan window.

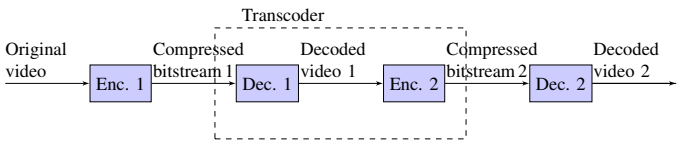



Fig. 16: Block diagram of transcoding process.

## V. Example application: image/video transcoding quality

Video transcoding consists of converting an already compressed video into a different format / using a different encoder (see Figure 16). Video transcoding is necessary in adaptive video streamingto create multiple representations of a video for content adaptation.

Calculating transcoding quality via objective quality metrics tends to be done with reference to the already encoded source image/video, but this does not provide information on the actual quality of the transcoded video with respect to the original uncompressed reference video [58], [59], [60], [61], [62], [63].

In the following, the assumptions and derivation above are used to derive an SSIM-based transcoding quality metric.

The error introduced by encoder 1 is denoted in the following as $e_1$, the error introduced by encoder 2 (transcoder) as $e_2$ and the total error as $e$. Focusing on the luminance component, composed of N pixels:

$$\mathcal{E}(e_1) = \frac{1}{N}\sum_{i=1}^{N} e_{1i} \tag{23}$$

$$\mathcal{E}(e_2) = \frac{1}{N}\sum_{i=1}^{N} e_{2i} \tag{24}$$

$$\mathcal{E}(e) = \frac{1}{N}\sum_{i=1}^{N} e_{i}. \tag{25}$$

It has been shown that the coding error of a DCT-based video codec has approximately zero mean (above and in [51]), i.e.:

$$\mathcal{E}(e_1) = \mathcal{E}(e_2) = 0 \tag{26}$$

$$\mathcal{E}(e) = 0. \tag{27}$$

For the variance:

$$\sigma_{e1}^2 = \mathcal{E}(e_1^2) - \mathcal{E}^2(e_1) = \mathcal{E}(e_1^2) = MSE_1 \tag{28}$$

$$\sigma_{e2}^2 = \mathcal{E}(e_2^2) - \mathcal{E}^2(e_2) = \mathcal{E}(e_2^2) = MSE_2 \tag{29}$$

$$\sigma_{e}^2 = \mathcal{E}(e^2) - \mathcal{E}^2(e) = \mathcal{E}(e^2) = MSE_{tot} \tag{30}$$

The variance of the total error can be written as:

$$\sigma_e^2 = \sigma_{e_1+e_2}^2 = \sigma_{e_1}^2 + \sigma_{e_2}^2 - 2\sigma_{e_1 e_2} \tag{31}$$

where the last term is the covariance.

With the assumption that $e_1$ and $e_2$ are independent:

$$\sigma_e^2 = \sigma_{e_1}^2 + \sigma_{e_2}^2 \tag{32}$$

hence:

$$MSE_{tot} = MSE_1 + MSE_2 \tag{33}$$

$$\begin{aligned} \mathrm{SSIM}_{tot} = \\ 1 - \frac{MSE_{tot}}{(2\sigma_y^2 + C_2)} = \\ 1 - \frac{MSE_1 + MSE_2}{(2\sigma_y^2 + C_2)} \end{aligned} \tag{34}$$

Considering that:

$$\mathrm{SSIM}_1 = 1 - \frac{MSE_1}{(2\sigma_y^2 + C_2)} \tag{35}$$

$$\mathrm{SSIM}_2 = 1 - \frac{MSE_2}{(2\sigma_y^2 + C_2)} \tag{36}$$

this can also be written as:

$$\begin{aligned} \mathrm{SSIM}_{tot} = 1 - \frac{MSE_1}{(2\sigma_y^2 + C_2)} - \frac{MSE_2}{(2\sigma_y^2 + C_2)} = \\ SSIM_1 + SSIM_2 - 1. \end{aligned} \tag{37}$$

We highlight that the quality of the transcoded video in terms of SSIM is neither the sum of the SSIM values of the two encoding steps, as considered by some authors, nor the mean of the SSIM values in the different steps, as considered by others. While it is quite obvious that either of the assumptions above are not correct, a relationship between the SSIM values before and after transcoding is not available in the literature according to the author's knowledge.

Similarly, in terms of PSNR:

$$SSIM_{tot} = 1 - \frac{\frac{(2^b-1)^2}{10^{PSNR_1/10}} + \frac{(2^b-1)^2}{10^{PSNR_2/10}}}{2\sigma_y^2 + C_2} \tag{38}$$

$$\begin{aligned} SSIM_{tot} = \\ 1 - \frac{(2^b-1)^2}{2\sigma_y^2 + C_2}\left(\frac{1}{10^{PSNR_1/10}} + \frac{1}{10^{PSNR_2/10}}\right) = \\ 1 - \frac{(2^b-1)^2}{2\sigma_y^2 + C_2}\left(10^{-PSNR_1/10} + 10^{-PSNR_2/10}\right) = \\ SSIM_1 + SSIM_2 - 1. \end{aligned} \tag{39}$$

$SSIM_1$ is the quality of the compressed video at the input of the transcoder. This value can be read from metadata when available or estimated in no-reference or reduced-reference modality based on bitrate and other information (in metadata / in the bitstream, such as motion vectors and QP values). An example method to predict SSIM with no reference to original video is provided in [43]. The method in [61] can also be used to estimate the objective quality metrics of the encoded mezzanine.

An example is reported in Figure 17. In this case the Baboon image is first encoded with JPEG with Quality Factor 70, then transcoded at multiple qualities, from QF 50 to QF 100. The global "ground truth" SSIM value (obtained comparing the transcoded version vs. original) is reported in black (with *). The SSIM value estimated via eq. (37) is reported in purple (dashed line). This is obtained calculating $SSIM_1$ comparing the original image vs. compressed at QF

70 and $SSIM_2$ comparing the transcoded image vs. the input version compressed at QF 70. In the first subfigure, a square window of size 16 has been considered for the SSIM evaluation, while in the second case the default Gaussian window has been used. The estimated global SSIM value follows very tightly the actual SSIM value in both cases, in particular in the second one. It was assumed here that a) the mean of the coding error in the coding process and the transcoder is zero and b) the coding errors produced by the two encoders are independent. Prediction errors are hence caused by possible inaccuracies in these assumption. In particular, the two types of coding errors may not be always independent. However, the proposed method provides a good approximation for the quality of transcoded images and video.

## VI. Conclusion

This paper highlights how the structural similarity metric SSIM for image quality assessment can be seen in many cases, such as DCT-based compressed images and video, as a content-aware version of the PSNR/MSE and SSIM can be obtained from PSNR/MSE via the variance of the impaired image/video, hence with no further reference to the original content. Different versions of simple and accurate SSIM estimations are proposed, that can be used alternatively depending on the context. This is expected to support the optimization of image/video compression and transmission systems, enabling mathematically tractable quality based optimization based on just one parameter beyond MSE/PSNR. The derivation also supports reduced-reference quality estimation based on SSIM and an easy comparison of different compression methods based on SSIM. In fact, BD-rate and BD-quality [64] were defined based on PSNR and their translation to different metrics such as SSIM is not obvious [65]. Finally, as an example application the assumptions and derivations above were used to show how the perceptual quality in terms of SSIM of a transcoded video can be calculated based on the SSIM values as a consequence of the first and second encoder, respectively.

## References


[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[2] Y. Wang, P. H. Chan, and V. Donzella, "Semantic-aware video compression for automotive cameras," *IEEE Transactions on Intelligent Vehicles*, 2023.

[3] R. Yan, Y. Liu, Y. Liu, L. Wang, R. Zhao, Y. Bai, and Z. Gui, "Image denoising for low-dose CT via convolutional dictionary learning and neural network," *IEEE Trans. on Computational Imaging*, vol. 9, pp. 83–93, 2023.

[4] B. Ding, R. Zhang, L. Xu, G. Liu, S. Yang, Y. Liu, and Q. Zhang, "$U^2D^2$Net: Unsupervised unified image dehazing and denoising network for single hazy image enhancement," *IEEE Transactions on Multimedia*, 2023.

[5] Y. Chen, J. Zhang, J. Zeng, W. Lai, X. Gui, and T.-X. Jiang, "A guidable nonlocal low-rank approximation model for hyperspectral image denoising," *Signal Processing*, p. 109266, 2023.

[6] L. Yan, M. Zhao, S. Liu, S. Shi, and J. Chen, "Cascaded transformer u-net for image restoration," *Signal Processing*, vol. 206, p. 108902, 2023.

[7] W. Li, J. Li, G. Gao, W. Deng, J. Zhou, J. Yang, and G.-J. Qi, "Cross-receptive focused inference network for lightweight image super-resolution," *IEEE Trans. on Multimedia*, vol. 26, pp. 864 – 877, 2023.

[8] N. Barman, Y. Reznik, and M. G. Martini, "A subjective dataset for multi-screen video streaming applications," in *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*, 2023, pp. 270–275.

[9] C. T. Hewage, A. Ahmad, T. Mallikarachchi, N. Barman, and M. G. Martini, "Measuring, modeling and integrating time-varying video quality in end-to-end multimedia service delivery: A review and open challenges," *IEEE Access*, vol. 10, pp. 60 267–60 293, 2022.

[10] J. Tang and G. Feng, "Background-detail restoration image deraining network based on convolutional dictionary network," *Signal Processing: Image Communication*, p. 117057, 2023.

[11] K. M. Nasr and M. G. Martini, "A visual quality evaluation method for telemedicine applications," *Signal Processing: Image Communication*, vol. 57, pp. 211–218, 2017.

[12] F. Russo, "A method based on piecewise linear models for accurate restoration of images corrupted by Gaussian noise," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 1935–1943, 2006.

[13] Q. Jiang, W. Gao, S. Wang, G. Yue, F. Shao, Y.-S. Ho, and S. Kwong, "Blind image quality measurement by exploiting high-order statistics with deep dictionary encoding network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7398–7410, 2020.

[14] S. Wang, D. Zheng, J. Zhao, W. J. Tam, and F. Speranza, "A digital watermarking and perceptual model based video quality measurement," in *2005 IEEE Instrumentation and Measurement Technology Conference Proceedings*, vol. 3. IEEE, 2005, pp. 1729–1734.

[15] R. Tu, G. Jiang, M. Yu, Y. Zhang, T. Luo, and Z. Zhu, "Pseudo-reference point cloud quality measurement based on joint 2D and 3D distortion description," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[16] T.-S. Ou, Y.-H. Huang, and H. H. Chen, "SSIM-based perceptual rate control for video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 5, pp. 682–691, 2011.

[17] L. Li, Z. Li, S. Liu, and H. Li, "Rate control for video-based point cloud compression," *IEEE Transactions on Image Processing*, vol. 29, pp. 6237–6250, 2020.

[18] C. Yeo, H. L. Tan, and Y. H. Tan, "On rate distortion optimization using SSIM," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 7, pp. 1170–1181, 2013.

[19] W. Gao, S. Kwong, Y. Zhou, and H. Yuan, "SSIM-based game theory approach for rate-distortion optimized intra frame CTU-level bit allocation," *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 988–999, 2016.

[20] T. Zhao, J. Wang, Z. Wang, and C. W. Chen, "SSIM-based coarse-grain scalable video coding," *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 210–221, 2015.

[21] A. Ichigaya and Y. Nishida, "Required bit rates analysis for a new broadcasting service using HEVC/H. 265," *IEEE Transactions on Broadcasting*, vol. 62, no. 2, pp. 417–425, 2016.

[22] Y. Li and X. Mou, "Joint optimization for SSIM-based CTU-level bit allocation and rate distortion optimization," *IEEE Transactions on Broadcasting*, vol. 67, no. 2, pp. 500–511, 2021.

[23] Q. Wang, S. Wang, M. Chen, and Y. Zhu, "A multiscale aligned video super-resolution network for improving vibration signal measurement accuracy," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[24] N. Barman, Y. Reznik, and M. Martini, "On the performance of video super-resolution algorithms for HTTP-based adaptive streaming applications," in *Applications of Digital Image Processing XLVI*, vol. 12674. SPIE, 2023, pp. 255–264.

[25] E. Shafiee and M. G. Martini, "Datasets for the quality assessment of light field imaging: Comparison and future directions," *IEEE Access*, vol. 11, pp. 15 014–15 029, 2023.

[26] K. Javidi, M. G. Martini, and P. A. Kara, "KULF-TT53: A display-specific turntable-based light field dataset for subjective quality assessment," *Electronics*, vol. 12, no. 23, p. 4868, 2023.

[27] A. Raj, A. K. Tiwari, and M. G. Martini, "Fundus image quality assessment: survey, challenges, and future scope," *IET Image Processing*, vol. 13, no. 8, pp. 1211–1224, 2019.

[28] N. Barman, Y. Reznik, and M. Martini, "Datasheet for subjective and objective quality assessment datasets," in *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*, 2023, pp. 111–114.

[29] S. Pezzulli, M. G. Martini, and N. Barman, "Estimation of quality scores from subjective tests – beyond subjects' MOS," *IEEE Transactions on Multimedia*, vol. 23, pp. 2505–2519, 2020.

[30] N. Barman and M. G. Martini, "User generated HDR gaming video streaming: dataset, codec comparison, and challenges," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1236–1249, 2021.

[31] Z. Li, C. Bampis, J. Novak, A. Aaron, K. Swanson, A. Moorthy, and J. Cock, "VMAF: The journey continues," *Netflix Technology Blog*, vol. 25, 2018.

[32] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.

[33] M. Martini, "A Simple Relationship Between SSIM and PSNR for DCT-Based Compressed Images and Video: SSIM as Content-Aware PSNR," in *2023 IEEE 25th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2023, pp. 1–5.

[34] A. De Angelis, A. Moschitta, F. Russo, and P. Carbone, "A vector approach for image quality assessment and some metrological considerations," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 1, pp. 14–25, 2008.

[35] T. Zinner, O. Abboud, O. Hohlfeld, T. Hossfeld, and P. Tran-Gia, "Towards QoE management for scalable video streaming," in *21th ITC specialist seminar on multimedia applications-traffic, performance and QoE*. Citeseer, 2010, pp. 64–69.

[36] A. R. Reibman and D. Poole, "Characterizing packet-loss impairments in compressed video," in *2007 IEEE International Conference on Image Processing*, vol. 5. IEEE, 2007, pp. V–77.

[37] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *2010 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 2366–2369.

[38] A. Horé and D. Ziou, "Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?" *IET Image Processing*, vol. 7, no. 1, pp. 12–24, 2013.

[39] R. Dosselmann and X. D. Yang, "A comprehensive assessment of the structural similarity index," *Signal, Image and Video Processing*, vol. 5, pp. 81–91, 2011.

[40] J. Silvestre-Blanes, "Structural similarity image quality reliability: Determining parameters and window size," *Signal Processing*, vol. 91, no. 4, pp. 1012–1020, 2011.

[41] J. Wang, P. Chen, N. Zheng, B. Chen, J. C. Principe, and F.-Y. Wang, "Associations between MSE and SSIM as cost functions in linear decomposition with application to bit allocation for sparse coding," *Neurocomputing*, vol. 422, pp. 139–149, 2021.

[42] G. Palubinskas, "Image similarity/distance measures: what is really behind MSE and SSIM?" *International Journal of Image and Data Fusion*, vol. 8, no. 1, pp. 32–53, 2017.

[43] T. Shanableh, "Prediction of structural similarity index of compressed video at a macroblock level," *IEEE Signal Processing Letters*, vol. 18, no. 5, pp. 335–338, 2011.

[44] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[45] S. S. Channappayya, A. C. Bovik, and R. W. Heath, "Rate bounds on SSIM index of quantized images," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1624–1639, 2008.

[46] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3378–3389, 2012.

[47] H. L. Tan, Z. Li, Y. H. Tan, S. Rahardja, and C. Yeo, "A perceptually relevant mse-based image quality metric," *IEEE Transactions on Image Processing*, vol. 22, no. 11, pp. 4447–4459, 2013.

[48] D. Brunet, E. R. Vrscay, and Z. Wang, "On the mathematical properties of the structural similarity index," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1488–1499, 2011.

[49] Y. Reznik, "Another look at SSIM image quality metric," *Electronic Imaging*, vol. 35, no. 8, pp. 305–1, 2023.

[50] J. Yang, G. Zhu, and Y.-Q. Shi, "Analyzing the effect of JPEG compression on local variance of image intensity," *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2647–2656, 2016.

[51] X. Shang, H. Zhao, G. Wang, X. Zhao, and Y. Zuo, "A novel objective quality assessment method for transcoded videos from H. 264/AVC to H. 265/HEVC utilizing probability theory," *IEEE Transactions on Broadcasting*, vol. 65, no. 4, pp. 777–781, 2019.

[52] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.

[53] H. Yuan, Q. Wang, Q. Liu, J. Huo, and P. Li, "Hybrid distortion-based rate-distortion optimization and rate control for H. 265/HEVC," *IEEE Transactions on Consumer Electronics*, vol. 67, no. 2, pp. 97–106, 2021.

[54] Y.-H. Huang, T.-S. Ou, P.-Y. Su, and H. H. Chen, "Perceptual rate-distortion optimization using structural similarity index as quality metric," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1614–1624, 2010.

[55] Y. Zhang, K. Ding, N. Li, H. Wang, X. Huang, and C.-C. J. Kuo, "Perceptually weighted rate distortion optimization for video-based point cloud compression," *IEEE Transactions on Image Processing*, vol. 32, pp. 5933–5947, 2023.

[56] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate-distortion optimization for video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 4, pp. 516–529, 2011.

[57] S. Al-Juboori and M. G. Martini, "Content characterization for live video compression optimization," in *13-th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, October 2024.

[58] S. Coulombe and S. Pigeon, "Low-complexity transcoding of JPEG images with near-optimal quality using a predictive quality factor and scaling parameters," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 712–721, 2009.

[59] V. V. Menon, R. Farahani, P. T. Rajendran, M. Ghanbari, H. Hellwagner, and C. Timmerer, "Transcoding quality prediction for adaptive video streaming," in *Proceedings of the 2nd Mile-High Video Conference*, 2023, pp. 103–109.

[60] S. M. A. H. Bukhari, K. Bilal, A. Erbad, A. Mohamed, and M. Guizani, "Video transcoding at the edge: cost and feasibility perspective," *Cluster Computing*, vol. 26, no. 1, pp. 157–180, 2023.

[61] R. Vanam, K. O. Lillevold, and Y. A. Reznik, "Assessing objective video quality in systems with multi-generation transcoding," in *Applications of Digital Image Processing XLIII*, vol. 11510. SPIE, 2020, pp. 201–207.

[62] J. Xu, M. Xu, Y. Wei, Z. Wang, and Z. Guan, "Fast H. 264 to HEVC transcoding: A deep learning method," *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp. 1633–1645, 2018.

[63] H. Yuan, C. Guo, J. Liu, X. Wang, and S. Kwong, "Motion-homogeneous-based fast transcoding method from H. 264/AVC to HEVC," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1416–1430, 2017.

[64] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *VCEG-M33*, 2001.

[65] N. Barman, M. G. Martini, and Y. Reznik, "Revisiting Bjontegaard delta bitrate (BD-BR) computation for codec compression efficiency comparison," in *Proceedings of the 1st Mile-High Video Conference*, 2022, pp. 113–114.


**Maria Martini** is a Professor (Chair) in the Faculty of Science, Engineering and Computing at Kingston University, London, UK, where she also leads the Wireless and Multimedia Networking Research Group. She received the Laurea in electronic engineering (summa cum laude) from the University of Perugia (Italy) in 1998 and the Ph.D. in Electronics and Computer Science from the University of Bologna (Italy) in 2002. She has led the KU team in a number of national and international research projects, funded by the European Commission (e.g., OPTIMIX, CONCERTO, QoE-NET, Qualinet), UK research councils (e.g., EPSRC, British Council, Royal Society), UK Technology Strategy Board / Innovate UK, and international industries. An IEEE Senior Member (since 2007) and Associate Editor for IEEE Signal Processing Magazine (2018-2020) and IEEE Transactions on Multimedia (2014-2018), she has been lead guest editor for the IEEE JSAC and guest editor for the IEEE Journal of Biomedical and Health Informatics, IEEE Multimedia, and the Int. Journal on Telemedicine and Applications, among others. She is part of international committees and expert groups, including the Video Quality Expert Group (VQEG) and the IEEE Multimedia Communications technical committee, where she has served as vice-chair (2014-2016) andchair (2012-2014) of the 3D Rendering, Processing, and Communications Interest Group. She is Expert Evaluator for the European Commission and EPSRC among others. She has authored over 200 scientific articles, contributions to international standardization groups (IEEE, ITU), and patents on wireless video. She chairs the IEEE P.3333.1.4 Standards Association Working Group.