Date and Time: Wednesday 23 June 2021 14:02:00 CEST

Job Number: 146784946

Documents (11)

1. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

2. Bibliography

3. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment1

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

4. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment2

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

5. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment3

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

6. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment4

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	**Narrowed by**
News	Source: IET Image Processing

7. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment5

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	**Narrowed by**
News	Source: IET Image Processing

8. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment6

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	**Narrowed by**
News	Source: IET Image Processing

9. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment7

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	**Narrowed by**
News	Source: IET Image Processing

10. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment8

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

11. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?_Attachment9

Client/Matter: -None-

Search Terms: HEADLINE("Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?")

Search Type: Terms and Connectors

Narrowed by:

Content Type	Narrowed by
News	Source: IET Image Processing

Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?

IET Image Processing

February 2013

Section: Pg. 12 - 24; Vol. 7; No. 1; 1751-9659

Length: 5084 words

Byline: AlainHoré, *alain.hore@usherbrooke.ca*

DjemelZiou

Body

1

Introduction

Any processing applied to an image may cause important loss of information or quality. Image quality refers to how the imaged scene is reproduced, and can be described by degradation in sharpness, dynamic range, colour blending, contrast, blocking effect or blur, for example. The ability to evaluate image quality is, nowadays, very important in several image processing applications such as art, remote sensing, e-commerce, information hiding and pattern recognition [1, 2]. Image quality evaluation methods can be subdivided into objective and subjective methods [3, 4]. Subjective methods are based on human judgment and operate on an informal basis, without reference to explicit criteria [5]. Objective methods are based on comparisons using explicit numerical criteria [6, 7], and several references are possible such as ground truth or prior knowledge expressed in terms of statistical parameters and tests [8, 9]. The mean-squared error (MSE), the peak signal-to-noise ratio (PSNR) [10] and the structural similarity index measure (SSIM) [11] are examples of quality measures, which are used in image analysis.

In this paper, we explain an analytical relationship between SSIM and PSNR for grey-level images. The PSNR is commonly used as a measure of the quality of reconstruction in image compression and image denoising [10]. Given a reference grey-level (8 bits) image f and a test image g, both of size $M \times N$, the PSNR between f and g is defined by (1)

$$\mathrm{PSNR}(f,\ g) = 10\log_{10}\left(255^2/\mathrm{MSE}(f,\ g)\right)$$

where (2)

$$\mathrm{MSE}(f,\ g) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij} - g_{ij}\right)^2$$

The PSNR value can be thought of as approaching infinity as the MSE approaches zero; this shows that a higher PSNR value implies a higher image quality. At the other end of the scale, a small value of the PSNR implies high numerical differences between images. SSIM was developed by Wang *et al*. [11] and is considered to be correlated with quality perception of the human visual system (HVS). Instead of using traditional error summation methods, the

SSIM is designed by modelling any image distortion as a combination of three factors that are loss of correlation, luminance distortion and contrast distortion. The SSIM is defined as (3)

$$\mathrm{SSIM}(f,\ g) = l(f,\ g)c(f,\ g)s(f,\ g)$$

where (4)

$$l(f,\ g) = \frac{2\mu_f \mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}, \quad c(f,\ g) = \frac{2\sigma_f \sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2}$$

$$s(f,\ g) = \frac{\sigma_{fg} + C_3}{\sigma_f \sigma_g + C_3}$$

We recall that: (μ_g and σ_g are computed similar to μ_f and σ_f by replacing f by g) (5)

$$\mu_f = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} f_{ij}}{MN}, \quad \sigma_f^2 = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(f_{ij} - \mu_f\right)^2}{MN}$$

$$\sigma_{fg} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(f_{ij} - \mu_f\right)\left(g_{ij} - \mu_g\right)}{MN}$$

The first term $l(f, g)$ in (4) is the luminance comparison function, which measures the closeness of the two images' mean luminance (μ_f and μ_g). The maximal value (i.e. equal to 1) of this factor is reached when $\mu_f = \mu_g$. The second term $c(f, g)$ is the contrast comparison function, which measures how the contrasts in the images are close to each other. Here, the contrast is measured by standard deviations σ_f and σ_g. This term is maximal (i.e. equal to 1) only if $\sigma_f = \sigma_g$. The third term $s(f, g)$ is the structure comparison function, which measures the correlation coefficient between the two images f and g. Note that σ_{fg} is the covariance between f and g. $s(f, g)$ is maximal (i.e. equal to 1) when $\sigma_{fg} = \sigma_f\sigma_g$. SSIM values are in [-1, 1], but we note that negative values are very rare in practise since they require images manipulated for this specific purpose. Thus in this paper, negative values of SSIM are irrelevant, and we use the interval [0, 1]. A value of 0 means no correlation with the original image, and 1 means that $f = g$. The positive constants C_1, C_2 and C_3 are used to avoid a null denominator. The authors propose in [11] to use $C_1 = (K_1L)2$, $C_2 = (K_2L)2$ and $C_3 = C_2/2$, where $K_1 = 0.01$, $K_2 = 0.03$ and $L = 255$ for grey-level images. We note that SSIM has been used extensively for different applications such as improving denoising and video-coding performances [12, 13], assessing video quality on digital TV [14] and there also exist image quality measures that are based on SSIM for improving the assessment of image quality [15, 16].

There are no precise rules for selecting SSIM or PSNR when the evaluation of image quality is required. Consequently, informal arguments and belief guide the interpretation of numerical values obtained during the evaluation process [17–21]. In fact, some studies have revealed that as opposed to SSIM, MSE (and hence PSNR) perform badly in discriminating structural content in images since various types of degradations applied to the same image can yield the same value of the MSE and thus the same value of the PSNR [22]. In some studies based on subjective human evaluation, it is reported that SSIM generally correlates better to human judgement than PSNR [23, 24]. Other studies have shown that MSE, and consequently PSNR, have the best performance in assessing the quality of noisy images [4]. In [9], in their attempt to compare some of the most common image quality measures in order to determine if there is any difference in their performance, the authors came to the conclusion that MSE, and consequently PSNR, can be related to SSIM. Those observations on the properties and performance of SSIM, PSNR and MSE are derived mainly from experimental studies. In [25], a relationship between MSE and the structure comparison function is derived as (6)

$$\mathrm{MSE}\left(\frac{f - \mu_f}{\sigma_f},\ \frac{g - \mu_g}{\sigma_g}\right) = 2\left(1 - s(f,\ g)\right)$$

This relationship is valid in the case where $C_3 = 0$ in (4). In [26], in their attempt to characterise packet-loss impairments in compressed video, the authors have indicated that SSIM can be estimated as follows using MSE (7)

$$\mathrm{SSIM}(f,\ g) = \frac{2\mu_f \mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \times \left[1 + \frac{\left(\mu_f - \mu_g\right)^2 - \mathrm{MSE}(f,\ g)}{\sigma_f^2 + \sigma_g^2 + C_2} \right]$$

This relationship is valid for $C_3 = C_2/2$ in (4). In this paper, we revisit the general formula of SSIM. First, we study the effect of the constants C_1, C_2 and C_3 on the values of SSIM. Second, we introduce an analytical relationship between SSIM and PSNR that can be used for predicting their similarity as well as their difference in evaluating image quality in the case of distortions such as Gaussian blur, additive Gaussian white noise, Jpeg and Jpeg2000 compression algorithms. Third, we show analytically that the PSNR is more sensitive than SSIM in discriminating images degraded using additive Gaussian white noise. This result is also confirmed experimentally. Note that we strictly focus on objective quantitative measurements in this paper, and we do not address any subjective or perceptual evaluation (although we give some references to previous works). The rest of the paper is organised as follows. In Section 2, we give a detailed description of the main objective of this paper, which is the derivation of an explicit analytical relationship between SSIM and PSNR. In Section 3, we make a series of tests using the Kodak images database, and we present some experimental results based on statistical models for comparing the two quality measures. The summary of our work with the final comments are given in Section 4.

2

Analytical relationship between PSNR and SSIM

To establish the relationship between SSIM and PSNR, we first derive the exactly analytical relationship between SSIM and MSE, and then we use that relationship to link SSIM to PSNR. This new exactly analytical relationship is further simplified through some assumptions that will be explained later. Note that in the remainder of the paper, the variables *f* and *g* are omitted in MSE, PSNR and SSIM. The MSE in (2) can be rewritten as (see the proof in Appendix 1) (8)

$$\mathrm{MSE} = \sigma_f^2 + \sigma_g^2 - 2\sigma_{fg} + \left(\mu_f - \mu_g\right)^2$$

where and

$$\sigma_g^2$$

 are the variance of images *f* and *g*, and σ_{fg} the covariance between *f* and *g*. The SSIM defined in (3) can be rewritten as (see the proof in Appendix 2) (9)

$$\frac{1}{\mathrm{SSIM}} = \frac{255^2 \times \alpha(f,\ g) \times e^{-\mathrm{PSNR} \times \ln(10)/10} + \beta(f,\ g)}{l(f,\ g)s(f,\ g)}$$

where (10)

$$\alpha(f,\ g) = \frac{1}{2\sigma_f \sigma_g + C_2}$$

$$\beta(f,\ g) = \frac{2\sigma_{fg} - \left(\mu_f - \mu_g\right)^2 + C_2}{2\sigma_f \sigma_g + C_2}$$

$$s(f,\ g) = \frac{\sigma_{fg} + C_3}{\sigma_f \sigma_g + C_3}$$

Equation (9) represents the complete and fundamental relationship between the SSIM and the PSNR that we derive, and it works for every image degradation or transformation for which the original and the degraded image are of the same size (note that the PSNR and the SSIM assume that the two images to compare have the same size). Even if it may appear a little complicated to analyse, we will see later that it can be simplified with some assumptions and for some particular image degradations. The constants C_2 and C_3 were introduced to avoid a null denominator [11]. Thus, in the case of non-null standard deviation values, the constants can be discarded. Non-null standard deviation values are found in real images on which at least one pixel has grey-level value different from the other pixels. For a better understanding of the constants C_2 and C_3, let us study how they affect the values of the contrast comparison function $c(f, g)$ and the structural comparison function $s(f, g)$ defined in (4). For this purpose, we consider a collection of original images that are degraded using 16 degradations classified into four groups which are additive Gaussian white noise, Gaussian blur, Jpeg and Jpeg200 compression algorithms (see Table 1). For each of these degradations, four parameters are applied, which enable us to take into consideration different levels or strengths of the degradations. Gaussian blur is the model of lens blurring and some atmospheric turbulence blurring [27]. Jpeg and Jpeg2000 compression algorithms are standards in imaging. Additive Gaussian noise is extensively used for assessing image-processing algorithms and for comparing quality measures. Moreover, sources of camera noise are generally considered as Gaussian [28]. In Table 1, we note that the noise standard deviations, denoted σ, are in the interval [0,1]. The grey levels of images have been normalised to [0,1] before adding noise. For each original and degraded image, we compute $c(f, g)$ for four values of C_2, which are 0, 1, 10 and 100. In the same way, we also compute $s(f, g)$ for four values of C_3 which are 0, 1, 10 and 100. To make the comparisons, we have used 76 original images of size 512 × 768 and 768 × 512 of the Kodak database, which are shown in Fig. 1, and we have also used blocks of size 64 × 64 and 16 × 16 within each image. Overall, almost 8 000 0000 computations of each of $c(f, g)$ and $s(f, g)$ have been performed in the experiments (we show in Appendix 3 computation of the exact number of tests performed). The luminance comparison function $l(f, g)$ is studied further in the paper. As the SSIM is defined for grey-level images, we first convert all the original images into grey-level images before applying the degradations and computing $c(f, g)$ and $s(f, g)$. In Fig. 2, we plot variation of the mean value of $c(f, g)$ and $s(f, g)$ for different sizes of images and different values of C_2 and C_3. The degradations are abbreviated as follows in Fig. 2: *G* is the Gaussian blur, *N* is the additive Gaussian white noise, *J* is the Jpeg compression and J2K is the Jpeg2000 compression. The abbreviations are followed by a number that indicates the corresponding parameter in Table 1. Note that the lines relating points in Fig. 2 do not represent continuity in the sense of continuous functions, but simply make more clear the differences between the mean values of $c(f, g)$ and $s(f, g)$ when varying C_2 and C_3. It appears that C_2 and C_3 make a greater impact on the values of $c(f, g)$ and $s(f, g)$ when the size of the images (or patches of images) becomes smaller. When the original images of size 512 × 768 and 768 × 512 are used, there is absolutely no change at all in the values of $c(f, g)$ and $s(f, g)$ as we vary the values C_2 and C_3. This can be explained by the fact that no image is completely uniform, meaning that the values of σ_f and σ_g are not null. In fact, σ_f and σ_g are high enough for those whole images to be sensitive to C_2 and C_3. When the images used are patches of size 64 × 64, we observe in Fig. 2 that there is absolutely no change in the values of $c(f, g)$ although there is some notable difference in the values of $s(f, g)$.

Fig. 1 Images of the Kodak database

Fig. 2 Variation of the mean value of c(f, g) and s(f, g) for different sizes of images and different values of C2 and C3

a and b Mean values of c(f, g) and s(f, g) as function of C2 and C3 for the 512 × 768 and 768 × 512 original images

c and d Mean value of c(f, g) and s(f, g) as function of C2 and C3 for image patches of size 64 × 64

e and f Mean value of c(f, g) and s(f, g) as function of C2 and C3 for image patches of size 16 × 16

Table 1 Degradations applied to the images with their corresponding parameters

Degradation	**Parameter 1**	**Parameter 2**	**Parameter 3**	**Parameter 4**
Gaussian blur (size of the filter)	3 × 3	5 × 5	7 × 7	9 × 9
Gaussian noise (noise variance)	?^{2} = 0.001	?^{2} = 0.01	?^{2} = 0.02	?^{2} = 0.05
	(? = 0.1)	(? = 0.14)	(? = 0.22)	(? = 0.03)

Degradation	Parameter 1	Parameter 2	Parameter 3	Parameter 4
Jpeg (quality ratio)	30%	50%	70%	90%
Jpeg2000 (quality ratio)	30%	50%	70%	90%

In the case of patches of size 16 × 16, we observe that the values of C_2 and C_3 affect the values of *c*(*f*, *g*) and *s*(*f*, *g*). In fact, this observation is not surprising since more patches of images may be completely uniform as the size of the patches become smaller, which means a null standard deviation. In other words, there is a greater probability that a patch of images is completely uniform as the size of the patch considered becomes smaller (like those found in the multi-scale SSIM, MSSIM). Consequently, the values of C_2 and C_3 will have in this case a non-negligible impact on *c*(*f*, *g*) and *s*(*f*, *g*). More interestingly, we can observe that for each degradation, the mean value of *c*(*f*, *g*) is monotonically increasing as a function of C_2, which means that the mean value of *c*(*f*, *g*) for $C_2 = a$ is greater than the mean value of *c*(*f*, *g*) for $C_2 = b < a$. In fact, this observation works not only for the mean value of *c*(*f*, *g*), but also for individual values of *c*(*f*, *g*) as we have noted in our experiments (almost 100% of all the images and patches of images tested with the various degradations). In Appendix 4, we analytically confirm this experimental observation. In a similar way, *s*(*f*, *g*) is monotonically increasing as a function of C_3. Consequently, choosing one value of C_2 and C_3 instead of other values may yield different values of SSIM: smaller values of C_2 and C_3 give smaller values of SSIM, while the opposite is observed for greater values of C_2 and C_3. However, because of the monotonic property of *c*(*f*, *g*) and *s*(*f*, *g*), the values of C_2 and C_3 are indeed not so much important if the SSIM is computed for comparing more than two images. When strictly comparing the quality of more than two images through SSIM, for example for assessing algorithms, we generally focus on determining which value of the SSIM is higher than the others; in this case, whatever the values of C_2 and C_3 considered for the comparisons, we get the same ordering of the SSIM values. Thus, we can discard the constants C_2 and C_3 for computing the SSIM. In the case of strictly comparing two images, whatever are the considered values of C_2 and C_3, we have nothing to do but to interpret or give a meaning to the single value of the SSIM so as to assess the similarity between the two images. Thus, even for that case which is a bit subjective, we can discard C_2 and C_3. In the general case of pairwise comparisons of image quality, for discarding C_2 and C_3, we assume that they are very small and satisfy

$$C_2 \ll \max(\sigma_f,\ \sigma_g) \text{ and } C_3 \ll \max(\sigma_f,\ \sigma_g)$$

; we can also simply set their values to zero. In such cases, we deduce from (9) and (10) that (the complete proof can be found in Appendix 5) (11)

$$\mathrm{PSNR} = -10\log_{10}\left[\frac{2\sigma_{fg}\,(l(f,\ g) - \mathrm{SSIM})}{255^2\mathrm{SSIM}} + \left(\frac{\mu_f - \mu_g}{255}\right)^2\right]$$

We recall that *l*(*f*, *g*) is defined in (4). The relationship described in (11) is general and can be further simplified for some image degradations. In fact, several tests were realised using the Kodak image database, which is shown in Fig. 1, have revealed that *l*(*f*, *g*) > 0.991

$$(\simeq 1)$$

for degradations because of Gaussian blur, additive Gaussian white zero mean noise, Jpeg and Jpeg2000 compression. This is summarised in Figs. 3*a*–*c* where we plot the mean value of *l*(*f*, *g*) for four sizes of images (512 × 768 and 768 × 512, 64 × 64 and 16 × 16) and four values of C_1 (0, 1, 10 and 100). Overall, almost 8 000 000 computations of *l*(*f*, *g*) have been performed (see Appendix 3 for more details on the number of tests performed). We can note that, for the degradations used, C_1 has no noticeable effect on the mean value of *l*(*f*, *g*), which means that it can simply be discarded. Moreover, the fact that *l*(*f*, *g*) is so close to 1 means that the luminance comparison function *l*(*f*, *g*) has a marginal and even negligible impact on the values of the SSIM defined in (3). Also, we observe that the shapes of the SSIM curves plotted in Figs. 3*d*–*f* are similar to those of *s*(*f*, *g*) shown in Figs. 2*b*, *d* and *f*. We can deduce that the structure comparison function *s*(*f*, *g*) has a greater impact on the values of SSIM than the contrast comparison function *c*(*f*, *g*). As *l*(*f*, *g*) is very close to 1, we can reasonably assume that *l*(*f*, *g*) = 1, which simply means that Gaussian blur, additive Gaussian white zero mean noise, Jpeg and Jpeg2000 compression algorithms all preserve the mean value of an image. Indeed, *l*(*f*, *g*) = 1 is equivalent to $\mu_f = \mu_g$. We presume here that the simplification *l*(*f*, *g*) = 1 does not depend at all on the original image *f* or the type of image, but on the type of degradation used. Consequently, we derive from (11) the following simplified relationship between PSNR and SSIM (12)

$$\mathrm{PSNR} = 10\log_{10}\left[\frac{255^2}{2\sigma_{fg}}\right] + 10\log_{10}\left[\frac{\mathrm{SSIM}}{1-\mathrm{SSIM}}\right]$$

Fig. 3 Variation of the mean value of l(f, g) and SSIM for different sizes of images

a–c Mean value of l(f, g) as function of C1 for original images of size 768 × 512 and 512 × 768, for patches of size 64 × 64 and for patches of size 16 × 16, respectively

d–f Mean value of SSIM as function of C1, C2 and C3 for original images of size 768 × 512 and 512 × 768, for patches of size 64 × 64 and for patches of size 16 × 16, respectively

As (12) indicates, there is an interesting link between PSNR and SSIM for the degradations used. It suggests that SSIM can be computed from PSNR and vice versa. This confirms the remarks of Dosselmann and Yang [9] who observed experimentally a possible link between MSE (and hence PSNR) and SSIM.

Fig. 4 is the plot of the PSNR as a function of SSIM, by varying σ_{fg} in the interval [0, 2552] in (12). It can be seen that all the curves have the same shape; they are equal up to an additive factor depending on σ_{fg}. In fact, σ_{fg} can be seen as the parameter that explains why some studies have revealed that different images with different structural degradations may yield the same value of MSE or PSNR while providing different values of SSIM [22]. As observed in Fig. 4, this is possible since for the same PSNR we can have different SSIM values corresponding to different σ_{fg} values. The converse is also true since the same value SSIM may correspond to different values of the PSNR. As the relationship between PSNR and SSIM is an objective function, we also state that for two different degradations, it may happen that the same value of SSIM corresponds to different values of MSE. In Fig. 5, the effect of the *log* term in (12) is presented. In fact, we can see that close values of the SSIM values produce different PSNR (and MSE) values (the contrast of the images has been increased to highlight the differences in a better manner). Here, we note that the difference is mostly visible for the MSE than the PSNR. In fact, since the PSNR is a logarithmic function, its variation is not always drastically notable, which is not the case for the MSE. In Fig. 6 (the contrast of the images has been increased for visualisation purposes), an example is shown where close values of SSIM produce non-negligible differences of the PSNR. Here also, the MSE produces the highest differences, which is not surprising. In conclusion, it appears that the PSNR and the MSE should no longer be considered as non-efficient image quality measures compared with SSIM on the basis that they may yield the same value for two degraded images whereas the SSIM yields different values. Indeed, the converse may also happen.

Fig. 4 Variation of PSNR as function of SSIM for different fixed values of σfg

Fig. 5 Example of very close SSIM values corresponding to different PSNR and MSE values

a Original image

b Jpeg, quality = 10%, SSIM = 0.977, PSNR = 29.69 and MSE = 69.84

c Gaussian noise, σ = 0.03, SSIM = 0.979, PSNR = 30.01 and MSE = 64.86

Fig. 6 Example of identical SSIM values corresponding to different PSNR and MSE values

a Original image

b Gaussian blur, size of the filter = 3 × 3, SSIM = 0.998, PSNR = 36.96 and MSE = 13.09

c Jpeg compression, quality = 70%, SSIM = 0.998, PSNR = 35.34 and MSE = 19.01

Another result that derived from Fig. 4 is that the PSNR is an increasing function with respect to SSIM. Although this can be easily observed experimentally, to the best of our knowledge, no analytical proof of this result has been reported in the literature. In fact, the derivative of the PSNR with respect to SSIM gives from (12) (13)

$$\mathrm{PSNR}' = \frac{\partial\,\mathrm{PSNR}}{\partial\,\mathrm{SSIM}} = \frac{10}{\ln(10)\mathrm{SSIM}(1-\mathrm{SSIM})}$$

Thus, we have PSNR′ > 0 in [0, 1], which confirms in this case that the PSNR is an increasing function with respect to SSIM. We note that when SSIM = 0 or SSIM = 1, PSNR′ is not defined and results in an infinite value.

Another interesting observation that can be made in Fig. 4 is that when SSIM is in the interval [0.2, 0.8], the curves are comparable with straight lines (an example is given by the red line plotted for the case σ_{fg} = 100). Computing the equation of the straight lines yields the approximated PSNR, denoted PSNR_{sl}, as follows in the interval [0.2, 0.8] (14)

$$\mathrm{PSNR}_{sl} = 20.069 \times \mathrm{SSIM} + \left(10\log_{10}\left(255^2/2\sigma_{fg}\right) - 10.034\right)$$

By plotting the absolute error (ΔP = PSNR - PSNR_{sl}) and the relative error (|ΔP|/PSNR) of the approximation as a function of SSIM in the interval [0.2, 0.8], we have observed that the maximum relative error is only 0.8%, which is small and indicates that linear approximation of the PSNR is accurate enough.

3

Objective study of the sensitivity of PSNR and SSIM to some image degradations

The relationship derived so far between SSIM and PSNR is quite interesting, but does not actually indicate whether or when one measure is more or less sensitive than the other according to a given image degradation. Thus, we need to evaluate the effects of degradations applied to images on the values of the PSNR and SSIM. For this purpose, comparisons of the PSNR and the SSIM values based on experiments using various original and degraded images are generally required [4, 9, 11, 19]. We use *F*-scores to measure the sensitivity of the PSNR and SSIM. More concretely, we measure how the PSNR and SSIM are influenced by the parameters of Gaussian white noise, Gaussian blur, Jpeg and Jpeg2000 compression, respectively, which were presented in Section 2. The images used for the experiments come from the Kodak database, shown in Fig. 1. To define the *F*-score, let us consider a set of parameter values of a given image degradation (for example, quality parameters of Jpeg compression = 30, 50, 70 and 90%). For each parameter, different values of the PSNR, forming a group, are computed for the original images. The same is made for SSIM. The *F*-score associated with the PSNR corresponds to the ratio of the variance of the set of mean values of the PSNR in all groups over the mean value of the within-group variances. The *F*-score of the SSIM is computed similarly. The *F*-score varies in [0, ∞[: a low value indicates that the parameters do not have a great impact on the values of the quality measure, meaning low sensitivity of the quality measure to the parameters; a high value of the *F*-score indicates the opposite (i.e. great impact and high sensitivity). Note that a similar approach was used in [4] to compare different quality measures. For research works on subjective evaluations of quality measures, one can refer to [23, 29].

In Fig. 7, we present the results of the *F*-score for various degradations in the form of a semi-logarithmic plot. As can been observed, SSIM appears to be more sensitive to Jpeg compression compared with the PSNR, whereas the opposite is observed for additive Gaussian white noise degradation. In fact, it is quite difficult to find a quality measure that is more sensitive to additive Gaussian noise than the PSNR, and some authors have noted that in their experiments [4].

Fig. 7 F-scores for different degradations

Still in Fig. 7, it appears that SSIM is slightly more sensitive than PSNR in discriminating the quality parameter of the Jpeg2000 compression, whereas the PSNR is slightly more sensitive than SSIM in discriminating the Gaussian blur. Finally, we note that the SSIM and PSNR are more sensitive to noise degradation than all the other degradations tested in this paper. Thus, the various structural distortions introduced by additive noise are the most distinguishable for both the PSNR and the SSIM compared with the distortions introduced by Gaussian blur, Jpeg and Jpeg2000 compression. In terms of applications, the PSNR should be preferred to the SSIM for denoising applications, while the SSIM should be preferred to the PSNR for applications that make use of Jpeg compression. For the applications that correspond to Gaussian blur or make use of Jpeg 2000 compression, either the PSNR or the SSIM can be used.

In Fig. 8, we present the box plots for the Jpeg compression algorithm and for additive Gaussian white noise degradation. In the box plots of the PSNR for example, each box corresponds to a group of computed values of the PSNR given a parameter value of a selected degradation. A graphical comparison based on box plots, where each box is centred on the group median and sized to the upper and lower 50 percentiles, allows one to see the distribution of the groups. The box plots are also related to the *F*-scores: if the *F*-score is high, there will be little overlap between two or more groups; if the *F*-score is not high, there will be a fair amount of overlap among all the groups. In Figs. 8*a*–*b*, the box plots of the PSNR and the SSIM for Jpeg compression show that there are more overlaps between the boxes in the case of the PSNR compared with the SSIM. This confirms that the PSNR is less sensitive than the SSIM to variation of the quality parameter values of the Jpeg compression algorithm. On the other hand, we note in Figs. 8*c* and *d*) that there is no overlap between the boxes of the PSNR in the case of additive Gaussian noise whereas some overlaps are visible for the SSIM. This explains to some extent the high value of the *F*-score of the PSNR observed in Fig. 7 for the additive Gaussian white noise. The flatness of the boxes of the PSNR in Figs. 8*c* and *d* and the lack of overlap between those boxes indicate a sharp sensitivity of the PSNR to variation of the variance of Gaussian noise. This sensitivity of the PSNR to noise is also reported in [4]: among twenty-six quality measures used to compare images (SSIM was not considered in that paper), the MSE, and so the PSNR, was the best to discriminate noise. More interestingly, it is possible to justify analytically the flatness of the boxes for the PSNR in Figs. 8*c* and *d*, and also to confirm that the PSNR is more sensitive to additive Gaussian noise than SSIM. The proof is as follows: let us consider an original image f and an image g obtained from f by adding a zero-mean Gaussian white noise n. We assume the practical case in which the original image and the noise are uncorrelated, meaning that the covariance $\sigma_{fn} = 0$. It is well known, and it can also be easily proved (from (1) and (8)), that the PSNR between f and $g = f + n$ is given by (15)

$$\mathrm{PSNR}(f,\ g = f + n) = 10 \log_{10}\left(\frac{255^2}{\sigma_n^2}\right)$$

where σ_n is the standard deviation of the noise. Also, by remembering that $l(f, g) = 1$ in the case of noisy images, it is possible to show that we get from (12) and (15) (16)

$$\mathrm{SSIM}(f,\ g = f + n) = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_n^2/2}$$

Fig. 8 Box plots for the Jpeg compression algorithm and for additive Gaussian white noise degradation

a and b Box plot of Jpeg compression for PSNR and SSIM, respectively

c and d Box plot of additive noise for PSNR and SSIM, respectively

By comparing (15) and (16), we see that the PSNR depends only on noise variance, whereas the SSIM depends both on noise variance and also original image (image variance). We note that it was also reported experimentally in [30] that SSIM does not depend entirely on noise variance. In fact, this result regarding the SSIM is not surprising since the SSIM has been designed to contain both the original image information and noise in order to account for visual masking (the SSIM accounts for HVS behaviour to some extent). The fact that the PSNR, in the case of additive Gaussian white noise, depends only on noise variance and not at all on the image used explains and justifies why the boxes of the PSNR in Figs. 8*c* and *d* are so flat: all images degraded using the same noise variance have the same PSNR value. Thus, the PSNR will always be more sensitive than SSIM for additive Gaussian white noise. In fact, it is quite difficult to find a quality measure that will be more sensitive to additive Gaussian white noise than PSNR. In Fig. 9, we present two images that were degraded using additive Gaussian noise with the same noise standard deviation $\sigma = 0.1$ (the pixel values were converted in the range [0,1] before adding noise). As we can observe, the PSNR is the same for the two degraded images, while the SSIM gives different values.

Fig. 9 Two images degraded with additive Gaussian white noise (σ = 0.1) yield the same PSNR value but different SSIM values

a Degradation with Gaussian noise, σ = 0.1, SSIM = 0.875, PSNR = 20.00 and MSE = 650.25

b Degradation with Gaussian noise, σ = 0.1, SSIM = 0.954, PSNR = 20.00 and MSE = 650.25

4

Conclusions

In this paper, we have compared the PSNR and the SSIM quality metrics by revisiting their analytical formula. The study reveals that there exists a simple logarithmic link between the PSNR and SSIM, which works for some common degradations such as Gaussian blur, additive Gaussian noise, Jpeg and Jpeg2000 compression. We have also undertaken an experimental study in order to assess the sensitivity of PSNR and SSIM to these degradations, that is, how the values of the parameter associated with each of these degradations affect the values of the PSNR and the SSIM. The study has revealed that the PSNR is more sensitive to additive Gaussian white noise than SSIM, whereas the opposite is observed for Jpeg compression. Both measures have slightly similar sensitivity to Gaussian blur and Jpeg2000 compression. In all cases, we have observed that the PSNR and the SSIM are more sensitive to additive Gaussian white noise than Gaussian blur, Jpeg and Jpeg2000 compression. We note that the analytical link derived between the PSNR and the SSIM explains and confirms various results that can be observed experimentally regarding these measures: for example, why two close values of the SSIM may yield two different values of the PSNR, and vice-versa.

As a final conclusion, it appears that the PSNR and the SSIM can be seen as related quality measures that mainly differ on their degree of sensitivity to image degradations.

Bibliography

REFERENCES

1 Yang C.H.: ‘Inverted pattern approach to improve image quality of information hiding by LSB substitution’, Pattern Recognit., 2008, 41, (8), pp. 2674–2683 (doi: 10.1016/j.patcog.2008.01.019)

2 Paglieroni D.W.: ‘Design considerations for image segmentation quality assessment measures’, Pattern Recognit., 2004, 37, (8), pp. 1607–1617 (doi: 10.1016/j.patcog.2004.01.017)

3 Kreis R.: ‘Issues of spectral quality in clinical H-magnetic resonance spectroscopy and a gallery of artifacts’, NMR Biomed., 2004, 17, (6), pp. 361–381 (doi: 10.1002/nbm.891)

4 Avcibas I., Sankur B., Sayood K.: ‘Statistical evaluation of image quality measures’, J. Electron. Imaging, 2002, 11, (2), pp. 206–223 (doi: 10.1117/1.1455011)

5 Farrell J.E.: ‘Image quality evaluation in colour imaging: vision and technology’, in MacDonald L.W., Luo M.R. (Eds.), (John Wiley, 1999), pp. 285–313

6 Cadik M., Slavik P.: ‘Evaluation of two principal approaches to objective image quality assessment’. Eighth Int. Conf. on Information Visualisation, 2004, pp. 513–551

7 Nguyen T.B., Ziou D.: ‘Contextual and non-contextual performance evaluation of edge detectors’, Pattern Recognit. Lett., 2000, 21, (9), pp. 805–816 (doi: 10.1016/S0167-8655(00)00045-3)

8 Medda A., DeBrunner V.: ‘Color image quality index based on the UIQI’. Proc. IEEE Southwest Symp. on Image Analysis and Interpretation, 2006, pp. 213–217

9 Dosselmann R., Yang X.D.: ‘Existing and emerging image quality metrics’. Proc. Canadian Conf. on Electrical and Computer Engineering, 2006, pp. 1906–1913

10 Chan L.C., Whiteman P.: ‘Hardware-constrained hybrid coding of video imagery’, IEEE Trans. Aerosp. Electron. Syst., 1983, 19, (1), pp. 71–84 (doi: 10.1109/TAES.1983.309421)

11 Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P.: ‘Image quality assessment: from error visibility to structural similarity’, IEEE Trans. Image Process., 2004, 13, (4), pp. 600–612 (doi: 10.1109/TIP.2003.819861)

12 Rehman A., Wang Z.: ‘SSIM-based non-local means image denoising’. IEEE Int. Conf. on Image Processing, 2011, pp. 217–220

13 Cui Z., Zui X.: ‘SSIM-based content adaptive frame skipping for low bit rate H.264 video coding’. Proc. IEEE Int. Conf. on Communication Technology, 2010, pp. 484–487

14 Rodriguez D.Z., Bressan G.: ‘Video quality assessments on digital TV and video streaming services using objective metrics’, IEEE Latin Am. Trans., 2012, 10, (1), pp. 1184–1189 (doi: 10.1109/TLA.2012.6142458)

15 Sylvestre-Blanes J.: ‘Structural similarity image quality reliability’, Elsevier Signal Process., 2011, 91, (4), pp. 1012–1020 (doi: 10.1016/j.sigpro.2010.10.003)

16 Chen G.H., Yang C.H., Po L.M., Xie S.L.: ‘Edge-based structural similarity for image quality assessment’. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, 2006, pp. II933–II936

17 Teo P.C., Heeger D.J.: ‘Perceptual image distortion’. Proc. First IEEE Int. Conf. on Image Processing, 1994, pp. 982–986

18 Van der Weken D., Nachtegael M., Kerre E.E.: ‘Image quality evaluation’. Proc. Sixth Int. Conf. on Signal Processing, 2002, vol. 1, pp. 711–714

19 Eskicioglu A.M., Fisher P.S.: ‘Image quality measures and their performance’, IEEE Trans. Commun., 1995, 43, (12), pp. 2959–2965 (doi: 10.1109/26.477498)

20 Sheikh H.R., Bovik A.C.: ‘Image information and visual quality’, IEEE Trans. Image Process., 2006, 15, (2), pp. 430–444 (doi: 10.1109/TIP.2005.859378)

21 Wang Z., Bovik A.C., Lu L.: ‘Why is image quality assessment so difficult?’ Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, 2002, pp. 3313–3316

22 Wang Z., Bovik A.C.: ‘Mean squared error: love it or leave it?’, IEEE Signal Process. Mag., 2009, 26, pp. 98–117 (doi: 10.1109/MSP.2008.930649)

23 Sheikh H.R., Sabir M.F., Bovik A.C.: ‘A statistical evaluation of recent full reference image quality assessment algorithms’, IEEE Trans. Image Process., 2006, 15, (11), pp. 3441–3452 (doi: 10.1109/TIP.2006.881959)

24 Ponomarenko N., Lukin V., Egiazarian K., Astola J., Carli M., Battisti F.: ‘Color image database for evaluation of image quality metrics’. Proc. IEEE 10th Workshop on Multimedia Signal Processing, 2008, pp. 403–408

25 Seshadrinathan K., Bovik A.C.: ‘Unifying analysis of full reference image quality assessment’. Proc. IEEE Int. Conf. on Image Processing, 2008, pp. 1200–1203

26 Reibman A.R., Poole D.: ‘Characterizing packet-loss impairments in compressed video’. Proc. IEEE Int. Conf. on Image Processing, 2007, pp. 2328–2331

27 Savakis A.E., Trussell H.J.: ‘Blur identification by residual spectral matching’, IEEE Trans. Image Process., 1993, 2, (2), pp. 141–151 (doi: 10.1109/83.217219)

28 Irie K., McKinnon A.E., Unsworth K., Woodhead I.M.: 'A technique for evaluation of CCD video-camera noise', IEEE Trans. Circuits Syst. Video Technol., 2008, 18, (2), pp. 280–284 (doi: 10.1109/TCSVT.2007.913972)

29 Gao X., Lu W., Tao D., Li X.: 'Image quality assessment based on multiscale geometric analysis', IEEE Trans. Image Process., 2009, 18, (7), pp. 1409–1423 (doi: 10.1109/TIP.2009.2018014)

30 Treibitz T., Schechner Y.Y.: 'Resolution loss without imaging blur', J. Opt. Soc. Am. A, 2012, 29, (8), pp. 1516–1528 (doi: 10.1364/JOSAA.29.001516)

Classification

Language: ENGLISH

Publication-Type: Magazine

Subject: ENGINEERING (73%); STATISTICAL METHOD (73%); TECHNICIANS & TECHNOLOGICAL WORKERS (73%); ELECTRONIC COMMERCE (56%); image denoising (%); Kodak database (%); SSIM (%); PSNR (%); image degradation (%); structural similarity index measure (%); peak-signal-to-noise ratio (%); image quality metrics (%)

Company: INFINITY AS (66%)

Industry: NAICS541512 COMPUTER SYSTEMS DESIGN SERVICES (66%); SIC7373 COMPUTER INTEGRATED SYSTEMS DESIGN (66%); COMPUTING & INFORMATION TECHNOLOGY (90%); IMAGE PROCESSING & COMPUTER VISION (90%); PROFESSIONAL SERVICES (90%); ENGINEERING (73%); ELECTRONIC COMMERCE (56%)

Load-Date: December 18, 2018

Bibliography

1. ***Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure***?, Alain Horé, Djemel Ziou, IET Image Processing, Pg. 12 - 24, (February 2013)

End of Document