ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/220931731

Image quality metrics: PSNR vs. SSIM

Conference Paper · August 2010

DOI: 10.1109/ICPR.2010.579 · Source: DBLP

CITATIONS
867

READS
21,676

2 authors:

Alain Horé
The University of Calgary
21 PUBLICATIONS **966** CITATIONS

SEE PROFILE

Djemel Ziou
Université de Sherbrooke
327 PUBLICATIONS **7,248** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Radon transform in computer vision View project

Project Blur images analysis and restoration View project

All content following this page was uploaded by Alain Horé on 24 March 2015.

The user has requested enhancement of the downloaded file.

Image quality metrics: PSNR vs. SSIM

Alain Horé
MOIVRE, Département d'informatique, Faculté des sciences, Université de Sherbrooke
Sherbrooke (Québec), Canada, J1K2R1
e-mail: alain.hore@usherbrooke.ca

Djemel Ziou
MOIVRE, Département d'informatique, Faculté des sciences, Université de Sherbrooke
Sherbrooke (Québec), Canada, J1K2R1
e-mail: djemel.ziou@usherbrooke.ca

***Abstract*—In this paper, we analyse two well-known objective image quality metrics, the peak-signal-to-noise ratio (PSNR) as well as the structural similarity index measure (SSIM), and we derive a simple mathematical relationship between them which works for various kinds of image degradations such as Gaussian blur, additive Gaussian white noise, jpeg and jpeg2000 compression. A series of tests realized on images extracted from the Kodak database gives a better understanding of the similarity and difference between the SSIM and the PSNR.**

Keywords-PSNR; SSIM; image quality metrics

I. INTRODUCTION

Any processing applied to an image may cause an important loss of information or quality. Image quality evaluation methods can be subdivided into objective and subjective methods [1, 2]. Subjective methods are based on human judgment and operate without reference to explicit criteria [3]. Objective methods are based on comparisons using explicit numerical criteria [4, 5], and several references are possible such as the ground truth or prior knowledge expressed in terms of statistical parameters and tests [6-8]. In this paper we explicit the relationship between the SSIM and the PSNR for grey-level (8 bits) images. Given a reference image f and a test image g, both of size $M \times N$, the PSNR between f and g is defined by:

$$PSNR(f,g) = 10\log_{10}\left(255^2 / MSE(f,g)\right) \quad (1)$$

where

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

The PSNR value approaches infinity as the MSE approaches zero; this shows that a higher PSNR value provides a higher image quality. At the other end of the scale, a small value of the PSNR implies high numerical differences between images. The SSIM is a well-known quality metric used to measure the similarity between two images. It was developed by Wang *et al.* [9], and is considered to be correlated with the quality perception of the human visual system (HVS). Instead of using traditional error summation methods, the SSIM is designed by modeling any image distortion as a combination of three factors that are loss of correlation, luminance distortion and contrast distortion. The SSIM is defined as:

$$SSIM(f,g) = l(f,g)c(f,g)s(f,g) \quad (3)$$

where

$$\begin{cases} l(f,g) = \dfrac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \\ c(f,g) = \dfrac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \\ s(f,g) = \dfrac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3} \end{cases} \quad (4)$$

The first term in (4) is the luminance comparison function which measures the closeness of the two images' mean luminance (μ_f and μ_g). This factor is maximal and equal to 1 only if $\mu_f = \mu_g$. The second term is the contrast comparison function which measures the closeness of the contrast of the two images. Here the contrast is measured by the standard deviation σ_f and σ_g. This term is maximal and equal to 1 only if $\sigma_f = \sigma_g$. The third term 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. The positive values of the SSIM index are in [0,1]. A value of 0 means no correlation between images, and 1 means that $f = g$. The positive constants C_1, C_2 and C_3 are used to avoid a null denominator.

There are no precise rules for selecting the SSIM or the 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 [10-13]. In fact, some studies have revealed that as opposed to the SSIM, the MSE and so the 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 [14]. Other studies have shown that the MSE, and consequently the PSNR, have the best performance in assessing the quality of noisy images [2]. The goal of this paper is to derive a simple analytical relationship between the SSIM and the PSNR that can be used to better understand their difference and similarity in the case of common degradations such as Gaussian blur, additive Gaussian noise, jpeg and jpeg2000 compression. We also compare in this paper the degree of sensitivity of the PSNR and the SSIM to those various degradations. In all of our study, we focus only on objective measurements and we do not address any subjective evaluation. The rest of the paper is organized as follows: in Section 2, we give a description of the derivation of a analytical relationship between the SSIM and the PSNR. In Section 3, we make a series of tests on natural images and we use some statistical models to compare the sensitivity of

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.579

IEEE computer society

the two quality measures to various degradations. We end the paper with concluding remarks.

II. ANALYTICAL RELATIONSHIP PSNR/SSIM

To establish the relationship between the SSIM and the PSNR, we first derive the relationship between the SSIM and the MSE, and then we use that relationship to link the SSIM to the PSNR. The MSE in equation (2) can be rewritten as:

$$MSE=\sigma_f^2+\sigma_g^2-2\sigma_{fg}+\left(\mu_f-\mu_g\right)^2 \tag{5}$$

where σ^2_f and σ^2_g are the variances of images f and g, and σ_{fg} the covariance between f and g:

$$\sigma_f^2=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij}-\mu_f\right)^2, \sigma_{fg}=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij}-\mu_f\right)\left(g_{ij}-\mu_g\right) \tag{6}$$

The SSIM defined in (2) can be rewritten as:

$$\frac{1}{SSIM}=\frac{255^2\times\alpha(f,g)\times e^{-PSNR\times\ln(10)/10}+\beta(f,g)}{l(f,g)s(f,g)} \tag{7}$$

where

$$\alpha(f,g)=\frac{1}{2\sigma_f\sigma_g+C_2}, \quad \beta(f,g)=\frac{2\sigma_{fg}-\left(\mu_f-\mu_g\right)^2+C_2}{2\sigma_f\sigma_g+C_2},$$
$$s(f,g)=\frac{\sigma_{fg}+C_3}{\sigma_f\sigma_g+C_3} \tag{8}$$

Let us now assume that $C_2 << \sigma_f, \sigma_g$ and $C_3 << \sigma_f, \sigma_g$. This assumption is made to nullify the effect of the constants appearing in the SSIM formula. We recall that these constants were introduced to avoid a null denominator [9]. 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 a grey-level value different from the other pixels. In such a case, we deduce from (7) and (8) that:

$$PSNR=10\log_{10}\left[\frac{2\sigma_{fg}\left(l(f,g)-SSIM\right)}{255^2\,SSIM}+\left(\frac{\mu_f-\mu_g}{255}\right)^2\right] \tag{9}$$

The relationship described in (9) is general and can be used for any kind of image degradation. This relationship can be further simplified in the case of some common image degradations. In fact, several tests realized using the Kodak image database, which is shown in Fig. 1, have revealed that $l(f,g)>0.991$ (≈ 1), for common and well known degradations such as Gaussian blur, additive Gaussian white noise, jpeg and jpeg2000 compression. All of these degradations generally introduce structural distortions of objects within images. The tests were realized by varying 16 parameters for each image: Gaussian blur (variances of the filter=0.5, 1, 1.5, 2), jpeg and jpeg2000 compression (quality parameters=30%, 50%, 70%, 90%), additive Gaussian white zero mean noise (standard deviation of the noise=0.03, 0.10, 0.14, 0.22: note that the noise standard deviations are in the interval [0,1] since the Matlab function used to generate noisy images begins by converting the grey levels into the interval [0,1] before adding noise). For each original and degraded image, we compute the luminance comparison function for three values of the constant C_1 which are 1, 10 and 100. To make the comparison, we have used the 76 images of size 512×768 and 768×512 extracted from the Kodak database as well as blocks of size 64×64 and 16×16 within each of these images. We note that all the images were first converted into grey-level images before the computations take place. In overall, almost 6 000 0000 computations of the luminance comparison function have been performed.

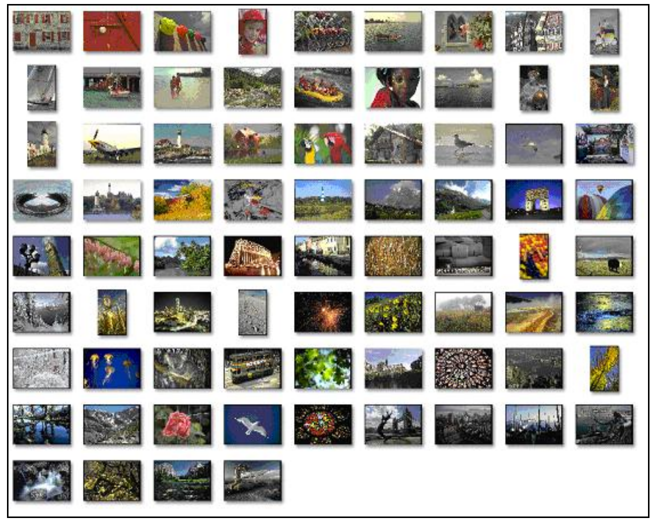

Figure 1. Images of the Kodak database

Using $l(f,g)=1$, which also means $\mu_f=\mu_g$, Equation (9) is rewritten as:

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

As Equation (10) indicates, there is an interesting link between the PSNR and the SSIM. It suggests that the values of the SSIM and those of the PSNR are not independent. This confirms the remarks of Dosselmann who noticed experimentally the existence of a possible link between the MSE (and so the PSNR) and the SSIM [8]. Fig. 2 is the plot of the PSNR as function of the SSIM, by varying σ_{fg} in the interval $]0,255^2]$ in Equation (10). It can be seen that all the curves have the same shape: they are equal up to an additive factor.

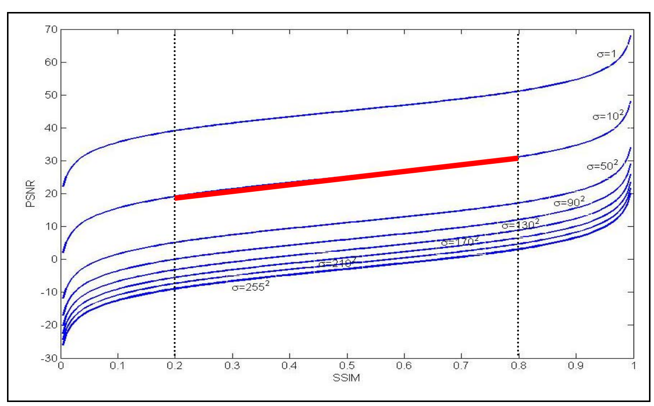

Figure 2. Variation of the PSNR as function of the SSIM for different fixed values of σ_{fg}

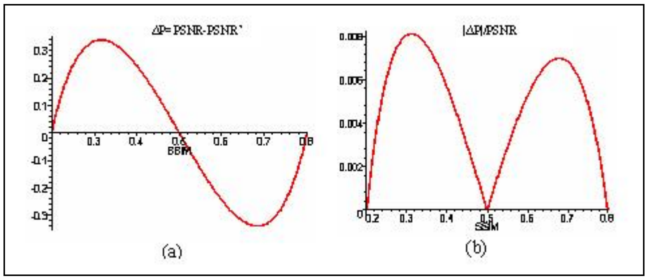

Figure 3. (a) Absolute error between the real and the approximated PSNR in the interval [0.2, 0.8]. (b) Relative error

Also, it appears in Fig. 2 that, when the SSIM varies in [0.2,0.8], the curves are essentially comparable to straight lines (an example is given by the red line plotted for the case $\sigma_{fg} = 10^2$). Computing the equation of the straight lines yields the approximated PSNR, denoted $PSNR_{sl}$, as follows in the interval [0.2,0.8]:

$$PSNR_{sl} = 20.069 \times SSIM + \left(10\log_{10}\left(255^2/2\sigma_{fg}\right)\text{-}10.034\right) \quad (11)$$

In Fig. 3, we plot the absolute error ($\Delta P = PSNR\text{-}PSNR_{sl}$) and the relative error ($|\Delta P|/PSNR$) of the approximation. As can been observed, the maximum relative error is only 0.8 %, which indicates that the linear approximation is accurate enough.

III. Experimental Results

The relationship derived so far between the SSIM and the PSNR is quite interesting, but does not actually indicate if one measure is more or less sensitive to any image degradation than the other. Thus, we have no information on how the values of the PSNR and the SSIM are influenced by any degradation applied to images. For this purpose, comparisons of the PSNR and the SSIM values based on experiments using various original and degraded images are generally required [2,8,9,12]. In this paper, we use F-scores to measure the sensitivity of the PSNR and the SSIM. More concretely, we measure how the PSNR and the SSIM are influenced by the parameters of the Gaussian 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%, 90%). For each parameter, different values of the PSNR, forming a group, are computed for the original images. The same is made for the SSIM. The F-score associated to 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 $[0,\infty[$: a low value indicates that the parameters do not have a great impact on the values of the quality measure, meaning a low sensitivity of the quality measure to the parameters; a high value of the F-score, on the contrary, indicates a great impact of the parameters on the values of the quality measure, meaning a high sensitivity. A similar approach was used in [2] to compare different quality measures.

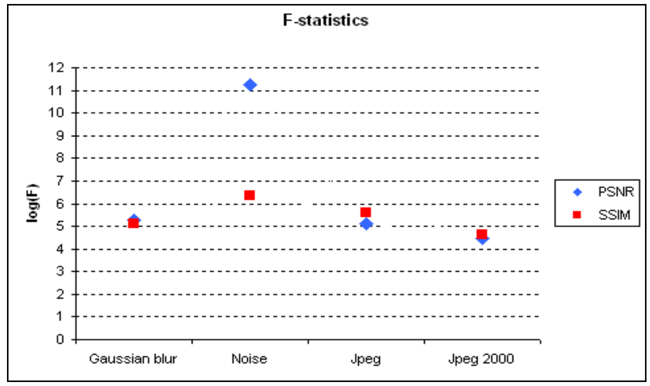

Figure 4. Comparison of the sensitivity of the PSNR and the SSIM using the F scores.

In Fig. 4, we present the results of the F-score for the various degradations. As can been observed, the SSIM seems to be more sensitive to jpeg compression compared to the PSNR, while the opposite is observed for additive Gaussian 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 noticed that in their experiments [2]. Still in Fig. 4, it appears that the SSIM is slightly more sensitive than the PSNR in discriminating the quality parameter of the jpeg2000 compression, while the PSNR is slightly better than the SSIM in discriminating the Gaussian blur. Finally, we note that the SSIM and the PSNR are more sensitive to noise degradation than all the other degradations tested in this paper. Thus, it appears that the various structural distortions introduced by additive noise are the most distinguishable for both the PSNR and the SSIM compared to the distortions introduced by Gaussian blur, jpeg and jpeg2000 compression.

IV. Conclusion

In this paper, we have undertaken a theoretical study to compare the PSNR and the SSIM quality metrics by analysing their analytical formula. The study has revealed that a simple analytical link exists between the PSNR and the SSIM, which works for 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 the PSNR and the SSIM to these degradations, that is how the values of the parameter associated to 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 noise than the SSIM, while 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 noise than Gaussian blur, jpeg and jpeg2000 compression.

As a final conclusion, it appears that the values of the PSNR can be predicted from the SSIM and vice-versa. The PSNR and the SSIM mainly differ on their degree of sensitivity to image degradations.

References

[1] R. Kreis, “Issues of spectral quality in clinical H-magnetic resonance spectroscopy and a gallery of artifacts”, NMR in Biomedecine, vol. 17, no. 6, pp. 361-381, 2004.

[2] I. Avcibas, B. Sankur and K. Sayood, “Statistical evaluation of image quality measures”, Journal of Electronic Imaging, vol. 11, no. 2, pp. 206-223, 2002

[3] J. E. Farrell, Image quality evaluation in colour imaging: vision and technology. MacDonald, L.W. and Luo, M.R. (Eds.), John Wiley, pp. 285-313, 1999.

[4] M. Cadik and P. Slavik, “Evaluation of two principal approaches to objective image quality assessment”, 8th International Conference on Information Visualisation, IEEE Computer Society Press, pp. 513-551, 2004.

[5] T. B. Nguyen and D. Ziou, “Contextual and non-contextual performance evaluation of edge detectors”, Pattern Recognition Letters, vol. 21, no.9, pp. 805-816, 2000.

[6] O. Elbadawy, M. R. El-Sakka, and M. S. Kamel, “An information theoretic image-quality measure”, Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering, vol. 1, pp. 169-172, 1998.

[7] A. Medda and V. DeBrunner, “Color image quality index based on the UIQI”, Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 213-217, 2006.

[8] R. Dosselmann and X. D. Yang, “Existing and emerging image quality metrics”, Proceedings of the Canadian Conference on Electrical and Computer Engineering, pp. 1906-1913, 2006.

[9] 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.

[10] P. C. Teo and D. J. Heeger, “Perceptual image distortion”, Proceedings of the 1st IEEE International Conference on Image Processing, pp. 982-986, 1994.

[11] D. Van der Weken, M. Nachtegael, and E. E. Kerre, “Image quality evaluation”, Proceedings of the 6th International Conference on Signal Processing, vol. 1, pp. 711-714, 2002.

[12] A. M. Eskicioglu and P. S. Fisher, “Image quality measures and their performance”, IEEE Transactions on Communications, vol. 43, no. 12, pp. 2959-2965, 1995.

[13] H. R. Sheikh and A. C. Bovik, “Image information and visual quality”, IEEE Transactions on Image Processing, vol. 15, no. 2, pp. 430-444, 2006.

[14] Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it?”, IEEE Signal Processing Magazine, vol. 26, pp.98-117, 2009.

View publication stats