SSIM Image Quality Metric for Denoised Images

PETER NDAJAH, HISAKAZU KIKUCHI, MASAHIRO YUKAWA,
HIDENORI WATANABE and SHOGO MURAMATSU
Department of Electrical and Electronics Engineering,
Niigata University,
JAPAN
email: ndajah@telecom0.eng.niigata-u.ac.jp

September 3, 2010

Abstract

The mean square error (MSE) and its related metrics such as peak signal to noise ratio (PSNR), root mean square error (RMSE), mean absolute error (MAE), and signal to noise ratio (SNR) have been the basis for mathematically defined image quality measurement for a long time. These methods are all based on the MSE. Denoisng quality has also been traditionally measured in terms of the MSE or its derivatives. But none of these metrics takes the structural fidelity of the image into account. Here, we investigate the structural changes that occur during the denoising process. In particular, we ascertain the structural fidelity of TV-denoised images.
Keywords: SSIM, MSE, TV, PSNR, NOISE, DENOISE, METRIC

1 MSE-based Image Quality Measure

The MSE has been the basis for image quality measure. Usually, one of the images (the original) is assumed to contain no distortions while the other image is contaminated by noise or some other kind of error. Suppose $\mathbf{x} = \{x_i|i = 1, 2, ..., N\}$ and $\mathbf{y} = \{y_i|i = 1, 2, ..., N\}$ where x_i and y_i are the ith samples in $\mathbf{x}$ and $\mathbf{y}$ and N is the number of signal samples. Then the MSE between the signals is

$$MSE(\mathbf{x},\mathbf{y}) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

$e_i = (x_i - y_i)$ is referred to as error signal. An image is a two dimensional signal so the MSE is given as

$$d(\mathbf{x},\mathbf{y}) = \left(\sum_{i=1}^{N}|e_i|^2\right)$$

Generally, in image processing, the MSE is often used in the form of the peak signal to noise ratio(PSNR) measure

$$PSNR = 10\log_{10}\frac{L^2}{MSE}$$

The PSNR is more useful than the MSE only when images of different dynamic ranges are being compared otherwise it is equivalent to the MSE [7].

2 Drawbacks of MSE-based Image Quality Measure

In imaging, the true aim of any denoising method is to improve the visual quality and fidelity of a noisy image but the MSE does not take into account image dependencies such as textures, orderings, patterns, etc. all of which affect image perception quality. Image pixel order transmit vital information about the structure of a visual scene. Unfortunately the MSE does not measure this. The correlation between the error signal and the underlying image significantly affects perceptual image distortion but this is also ignored by the MSE. The MSE does not take into account the signs of the error (since its square is used) signal added to an image. However, the visual fidelity of the resulting image has been proved to be drastically different. Since all images are treated equally in the formulation of the MSE, image content-dependent variations in image fidelity cannot be accouted for.

3 Structural Similarity (SSIM)

The SSIM is a recently proposed image fidelity measure which has proved highly effective in measuring the fidelity of signals. The SSIM approach was originally motivated by the observation that natural images have highly structured signals with strong neighborhood dependencies. These dependencies carry useful information about the structures of the objects in the visual scene.

The human visual system is highly adapted to extract structural information from visual scenes. For this reason, image fidelity measurement should retain the signal structure as an important content. A distinction has to be made between non-structural distortions such as variations in luminance, contrast, Gamma distortions, and spatial shift(these do not change the structure of the image in any way) and the structural distortions such as additive Gaussian noise, blur and lossy compression(e.g. JPEG). These distort the structure of the image significantly.

The human visual system is highly sensitive to structural distortions and easily compensates for non-structural distortions. The main function of the SSIM is to simulate this functionality.

Let $\mathbf{x} = \{x_i | i = 1, 2, ..., N\}$ and $\mathbf{y} = \{y_i | i = 1, 2, ..., N\}$ be the original and the test image signals respectively. Then, the SSIM

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]} \quad (1)$$

The above equation can be rewritten as

$$Q = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (2)$$

The SSIM measures distortions as a combination of three factors: loss of correlation, luminance distortion and contrast distortion. The first component in (2) is the correlation coefficient between $\mathbf{x}$ and $\mathbf{y}$. It measures the degree of correlation between $\mathbf{x}$ and $\mathbf{y}$. Its dynamic range is $[-1, 1]$ and the best value 1 is obtained when y_i is linear with respect to x_i for all $i = 1, 2, ..., N$ i.e. $y_i = ax_i + b$. The second component has a value range of $[0, 1]$. It measures the mean luminance between $\mathbf{x}$. It equals 1 if and only if $\bar{x} = \bar{y}$. The third compnent meaures the similarity of the contrast between $\mathbf{x}$ and $\mathbf{y}$. Its range is also $[0, 1]$, where the best value is 1. This occurs only when $\sigma_x = \sigma_y$.

4 A Comparison of the MSE and the SSIM

Figure 1: Images with different structural distortions but the same MSE values

Figure 1 illustrates the shortcomings of the MSE. In all the images shown, the MSE $= 255$ even when the visual structures are greatly distorted. The SSIM on the other hand seems to reflects the structural changes in the images more faithfully. This is the advantage of the SSIM over the MSE. The human visual system (HVS) is very sensitive to structural changes, therefore any metric that will be well correlated to the HVS must take into account the structural dependencies of the signal samples in order to provide effective pre-

ISSN: 1792-6130

ISBN: 978-960-474-246-2

Figure 2: Denoised images showing values MSE values and SSIM index values

dictions of image quality. As often happens during denoising of images, structural changes such as blurring can happen. Most denoising algorithms do not actually 'remove' the noise. It is more a process of noise minimization rather than removal. The amount of noise still left in the image sample after the denoising operation depends on the amount of noise originally in the image before the denoising operation. But the MSE-based metrics may not be able to capture this reality because they are not designed to to measure the structural distortions that may occur.

5 Denoised Image Structural Fidelity

So why use the SSIM index to measure the quality of denoised images? Because the MSE-based metrics do not tell the whole story. The ultimate objective of denoising is to produce an image that is judged to be a good representation of the reference image (known or unknown). The HVS is the ultimate judge of what a good quality image is. This means that the structural fidelity of the denoised image is of utmost importance because the HVS uses the structural fidelity to measure the quality of an image. The MSE-based metrics fail to measure the structural improvement or degradation in an image after denoising. This is because in the MSE-based metrics, the signal samples are considered to be independent of each other. As we can see in Figure 2, the denoised images have different SSIM values (as judged by the HVS) but they have practically the same MSE values.

The total variation denoising algorithm was used to denoise the images because of its effectiveness and also because it has tunable parameters λ and τ that control the effectiveness of the denoising process. We have varied the values of λ and kept τ constant in the experiments.

6 Conclusion

We used the lena image as the test image in our experiments. As Figure 2 shows, the changes in structural similarity indices of the images correlate somewhat with human visual system. For example, when $\lambda \leq 2$, ((d)-(f)), the algorithm causes blurring in the images. The SSIM index reflects this fact as the SSIM values become progressively smaller with reducing visual quality of the images, However, the MSE remained the same throughout our experiments. for this reason, it may be useful to use the SSIM as an alternative metric of denoised image quality since it is a good measure of the structural degradation or improvement in a denoised image.

References

[1] Z. Wang, L. Lu and A.C. Bovik. Video Quality Assessment based on Structural Distortion Analysis. *IEEE International Conference on Image Processing,* Genoa, Italy, Sept. 11-14, 2005.

[2] Z. Wang, A.C. Bovik and E.P. Simoncelli. *Structural Approaches to Image Quality Assessment* in *Handbook of Image and Video Processing* (Al Bovik, ed.), 2nd Edition, Academic Press, 2005.

[3] Z. Wang, A.C. Bovik and L. Lu. Why is Image Quality Assessment Difficult? *IEEE International Conference on Acoustics, Speech and Signal Processing,* May, 2002

[4] Z. Wang and A.C. Bovik. A Universal Image Quality Index. *IEEE Signal Processing Letters,* vol. 9. no.3, pp. 81-84, March 2002

[5] 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

[6] Zhou Wang and Alan C. Bovik. Mean Squared Error: Love it ot Leave it? *IEEE Signal Processing Magazine,* January 2009.

[7] Zhou Wang and Alan C. Bovik. A Universal Image Quality Index. *IEEE Signal Processing Letters,* vol. 9 no. 3, pp. 81-84, 2002.