A Processing Algorithm with some User-Defined Parameters

Jeewon Hwang and Juphil Cho*

*Dept. of Information Technology, CAIT, Chonbuk National University, Jeonju, Korea
Dept. of Radio Communication Engineering, Kunsan National University, Kunsan, Korea*

*hwangj@jbnu.ac.kr, *Stefano@kunsan.ac.kr (Corresponding author)*

Abstract

A simple processing algorithm with a user-controlled parameter is presented. The average value of locally computed correlation coefficients between the two processed images is found. This criterion is closely related to the way in which human observer assesses image quality. Finally, image characteristic evaluation is computed as the average value of locally computed correlation coefficients, adjusted by the average correlation coefficient between the reference and error images. By this approach the proposed evaluation differentiates between the random and signal dependent distortions, which have different effects on human observer. Performance of the proposed characteristic evaluation is illustrated by examples involving images with different types of degradation

Keywords: Contrast sensitivity function, Image noise, Image processing

1. Introduction

Objective (quantitative) measures use intensities from the distorted input image and the reference image to compute a value that indicates image quality. The goal is to find a good measure which produces results which are in good agreement with subjective evaluation. Most widely used objective measures are mean squared error (MSE) and MSE-based measures: peak signal to noise ratio (PSNR) and signal to noise ratio (SNR). These simple measures work well when images with the same type of distortion or degradation are compared. In this case distorted image with smaller MSE will be perceived closer to the original image than the one with greater MSE. However, when images with different types of distortion (additive white Gaussian noise, blurring. contrast stretching, JPEG2000 coding/compression) are compared MSE does not produce results that correlate well with subjective quality assessment. Images with different types of distortions with the same MSE values can have very different subjective visual qualities. One role of these measures is to assess performance of algorithms for image enhancement and coding. Another role is to optimize image processing algorithms. For example, in image coding the goal is to represent images with as few bits as possible with the minimal loss of image quality. Quality measure is used as a criterion in a sense that the bits that contribute more significantly to the used quality measure are coded first. The objective video quality measures have been described recently in [1]. Examples of digital video distortions including tiling, smearing, jerkiness, edge busyness, and object retention that are produced by video processing systems are fundamentally different kinds of impairments than those produced by image processing systems. So both subjective and objective quality

assessments are quite different for images and videos. In this paper we focus our attention only on image quality monitoring and assessment.

These algorithms try to assess image quality by taking into account the properties of human visual system (HVS). Most of these algorithms are trying to model the following three properties of HVS: nonlinear relationship between image intensities and perceived brightness, frequency response of HVS (contrast sensitivity function) and texture masking [1-2]. The algorithm described in [2] can identify the pixels with visible distortion, but it cannot quantify the extent of easily visible distortion, because every spatial location with clearly visible distortion will have detection probability of one in the output map. Also, this algorithm does not produce one number that indicates image quality. In [2], the properties of HVS are used to design an algorithm for image coding. In this approach input image is first decomposed into 16 sub-bands using generalized quadrature mirror filters. Image quality measure in this case is defined as MSE of normalized sub-band coefficients of the original image and distorted image, where each sub-band coefficient is normalized by the perceptual threshold for that sub-band at that spatial location. Another approach presented in [2], assumes that input image is degraded by two sources of degradation: linear distortion (for example as a result of filtering) and additive noise. In the first step linear distortion and additive noise are separated and then two quality measures are computed: distortion and noise quality measures. Distortion measure is computed using discrete Fourier transform (DFT) coefficients of the original image and original image distorted by linear distortion only. Noise quality measure is computed using the original image degraded by linear distortion and noise and the original image degraded by linear distortion only. This approach requires as input not only two images but also the algorithm that causes linear distortion and produces two numbers, which assess image quality.

Problems with the types of image characteristic evaluation similar to those described in [1-3] are that there are no standards for modeling the same properties of HVS. For example, models for brightness perception in [1] and [2] are totally different. Also, there are various contrast sensitivity functions estimated by different authors. Another problem is computation of perceptual threshold, which different authors compute in a totally different way. This sometimes requires computation of local contrast (as in [1] and [3]), which is difficult to define for complex images. A different image characteristic evaluation called the universal quality index (UQI) was proposed in [5] and [6]. This measure is computed locally to model any distortion and it is defined as a product of three components: correlation, mean and variance of distortion. Most important of these components is the correlation coefficient, which measures the degree of linear relationship between the corresponding blocks of pixels. Characteristic evaluation for the whole image is the average value of locally computed values. Most other approaches try to modify MSE by modeling properties of HVS. This approach defines different criterion which is better related to the way in which human observer assesses image quality. No HVS model is used here.

The approach presented here will use the basic idea from [5-7], but it will not use image characteristic evaluation as it is defined there (only correlation coefficient will be used). In the first step, the reference and distorted images are processed by a simple model of HVS. It consists of a nonlinear function modeling the brightness perception and a 2-D filter modeling the frequency response of HVS (contrast sensitivity function). The 2-D filter used in this HVS model is not fixed. It contains one user-defined parameter which can be changed depending on the content of the reference image. After this, the correlation coefficient is computed on a block-by-block basis for the processed

input images. Finally, the characteristic evaluation is computed as the average correlation coefficient between the reference image and distorted image, adjusted by the average correlation coefficient between the reference image and error image. This is a simpler model to those described in [5-9].

2. A Simple Image Characteristic Evaluation

Perceived brightness is a nonlinear function of intensity and this is modeled by transforming input intensities by some monotonically increasing nonlinear function. Various functions have been suggested to model this. One frequently cited result states that brightness perceived by HVS is proportional to the logarithm of intensity as it is known as Weber's law [8]. HVS model developed in [15] uses cube root function to model this effect. Some use brightness perception models developed for specific display device and an example of this can be found in [6]. Logarithmic or cube root functions are not accurate models for brightness perception by HVS especially at low intensities. According to them the sensitivity to intensity increases is greatest at low intensities and decreases at mid and high intensities. This is clearly wrong, because an increase in intensity result in a very small increase in perceived brightness, when the intensity is low. The following nonlinear function will be used here instead [12].

$$B = \begin{cases} 0 & \text{if } 0 \le I \le I_{TR} \\ \dfrac{B_{MAX}}{2}\left(\dfrac{2(I - I_{TR})}{I_{MAX} - I_{TR}}\right)^2 & \text{if } I_{TR} \le I < (I_{MAX} + I_{TR})/2 \\ B_{MAX} - \dfrac{B_{MAX}}{2}\left(\dfrac{2(I_{MAX} - I)}{I_{MAX} - I_{TR}}\right)^2 & \text{if } (I_{MAX} + I_{TR})/2 < I \le I_{MAX} \end{cases} \tag{1}$$

Where, $I_{TR} = 20$ is threshold value for intensity, $I_{MAX} = 255$ is the maximum value of intensity and $B_{MAX} = 100$ is the maximum value of perceived brightness. When the intensity goes from 0 to 20 there is no increase in perceived brightness and when intensity goes from 20 to 255, perceived brightness goes from 0 to 100 as shown in Figure 1. This range is chosen for convenience. The role of this function is to emphasize intensities in the mid range and de-emphasize very high and very low intensities, which approximates the brightness perception by HVS.

The second component of HVS model is a 2-D filter, which models frequency response of HVS. HVS is not equally sensitive to all spatial frequencies and it is a frequency selective system. If HVS were not a frequency selective system, then we would have seen the same amount of noise in each image. Since we see less noise at very low and very high frequencies than in the mid frequency range, HVS must be a band pass system. This is modeled by a contrast sensitivity function (CSF), which represents the frequency response of HVS. CSF models HVS as a 2-D filter. Spatial frequency is expressed in cycles per unit visual angle (degree) instead of cycles per unit length (cm). This is because images can be observed form the various distances and the same spatial frequency expressed in cycles/cm will be perceived differently for different viewing distances. Various functions have been suggested to model this effect. Most are band pass in nature but maximum at different frequencies. For example, it was noted in [9, 10] that various functions suggested by different authors have maximum at frequencies ranging from 3 to 10 cycles/degree. HVS model presented in [12] uses a high pass filter. The reasoning behind this is that HVS treats near and far objects in the same way.

Since there is no reliable way to tell which one of these functions represents the best model of HVS, none of them will be used here. Instead, CSF given by the following formula will be used [12]:

$$H(f) = \begin{cases} (0.0512 + 0.8512 f) \exp(-(0.3192 f)), \\ \quad \text{for } f \leq 3 \, \text{cycles/degee} \\ 1, \quad \text{for } 3 \, \text{cycles/degee} < f < f_0 \\ \exp(-0.1(f - f_0)^{1.1}), \quad \text{for } f > f_0 \end{cases} \tag{2}$$

Where, f denotes frequency in cycles/degree and f_0 represents the frequency in cycles/degree at which CSF starts to decrease exponentially. This parameter is user-defined and depends on the reference image. If the reference image contains one large object (for example "Lena" image) than f_0 takes lower values than in the case when the reference image contains smaller objects or lot of fine details. Choice of values for this parameter will be discussed in Section 3. This requires some user intervention but it yields better results. This function is shown in Figure 1 for f_0=5 cycles/degree.
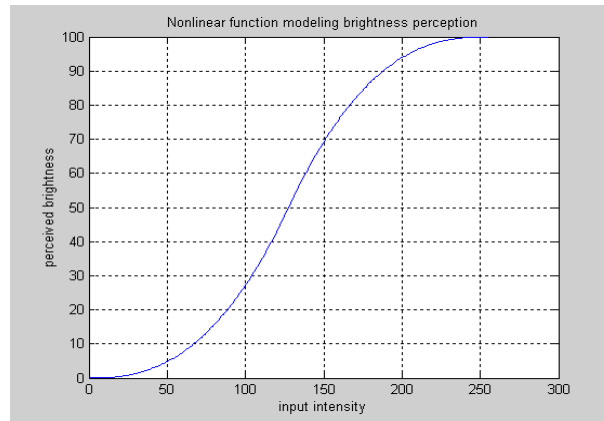

Figure 1. Nonlinear modeling of brightness perception

MSE and MSE-based measures such as PSNR and SNR are widely used to assess performances of image enhancement and coding algorithms. MSE-based measures work well when images with the same type of degradation are compared. In a sequence of images obtained by distorting an image by various amounts of the same type of degradation, the image with the smallest MSE will be perceived by a human observer to be the closest to the original image. However, when images with different types of degradation are compared, the image with the smallest MSE will not always be perceived to be the closest to the original image. It is possible to create images with the same MSE value but with totally different visual quality.

Digital images are represented using a finite number of intensity levels. For example, gray scale images with 8 bits/pixel are represented by 256 intensity levels, where 0 corresponds to the darkest level and 255 corresponds to the brightest level. Intensities between these two extremes represent various shades of gray level. Perceived brightness is a nonlinear function of intensity and this is modeled by transforming input intensities by some monotonically increasing nonlinear function. Various functions have been suggested to model this.

The original or reference image is denoted by f(m,n) and the distorted image is denoted by p(m,n). The size of both these images is M x N pixels. The processing of the input images f(m,n) and p(m,n) by this model will produce images x(m,n) and y(m,n) respectively, which will be used to evaluate the characteristic evaluation. Processed images x(m,n) and y(m,n) can be used to compute weighted MSE (WMSE) as was done in [4]. Instead of this, a different procedure, similar to what was done in [2] and [3], will be used here. The processed images x(m,n) and y(m,n) are partitioned into 8x8 pixel blocks. For each of these blocks, the correlation coefficient between the corresponding samples from images x(m,n) and y(m,n) is computed. If we denote the correlation coefficient value between x(m,n) and y(m,n) in the i-th block of data by $\rho_{xy}(i)$, then we can find the average correlation coefficient value between the x(m,n) and y(m,n) as:

$$\rho_{xy\_avg} = \frac{1}{L}\sum_{i}^{L}\rho_{xy}(i) \qquad (2)$$

where, L is the total number of data blocks and the correlation coefficient between two vectors $X = [x_1 \ x_2 \ \dots \ x_N]$ and $Y = [y_1 \ y_2 \ \dots \ y_N]$ is defined as $\rho_{XY} = \sigma_{XY} / \sigma_X \sigma_Y$ where σ_{XY} is the covariance and σ_x, σ_y are the standard deviations of X and Y.

Then local correlation coefficients are computed block by block between images x(m,n) and e(m,n). If the local correlation coefficient for the ith block of data is denoted by $\rho_{xe}(i)$, then the average correlation coefficient between images x(m,n) and e(m,n) is:

$$\rho_{xe\_avg} = \frac{1}{L}\sum_{i}^{L}\rho_{xe}(i) \qquad (3)$$

This coefficient will have very small positive or negative value close to zero, when there is little or no correlation between the images x(m,n) and e(m,n) and this will indicate random noise. On the other hand if there is a significant correlation between x(m,n) and e(m,n) the average correlation coefficient ρ_{xe_avg} will have a higher absolute value, indicating signal dependent noise. Finally, the proposed new image characteristic evaluation denoted by Q is defined as:

$$Q = sign\{\rho_{xy\_avg}\}\,|\,\rho_{xy\_avg}\,|^{f(\rho_{xe\_avg})}$$

where   (4)

$$f(\rho_{xe\_avg}) = 1.2 + 0.5\tanh\left(\frac{|\,\rho_{xe\_avg}\,|-0.3}{0.15}\right)$$

The characteristic evaluation (Q) defined by equation (4) needs some explanation. It will have the same sign as ρ_{xy_avg} given by equation (2), but its magnitude will be modified value of the magnitude of ρ_{xe_avg}. This modification is made according to the way the average correlation coefficient changes. If the magnitude of ρ_{xe_avg} is small (close to zero), then it indicates random noise. In brief, the proposed scheme can be summarized in the following four steps:

Given the reference image f(m,n), distorted image p(m,n), width, height and number of pixels of the input images and viewing distance, compute images x(m,n) and y(m,n) using the described model of HVS.

Compute ρ_{xy-avg} as the average value of locally computed correlation coefficients between images x(m,n) and y(m,n).

Compute ρ_{xe_avg} as the average value of the locally computed correlation coefficients between images x(m,n) and e(m,n)

Find the image characteristic evaluation Q using equation (4).

3. Simulation Results

Performances of the algorithm will be illustrated by several examples using the following parameter values: equal image height and width, viewing distance equal to four times image height and image size of 512 x 512 pixels. Parameter f_0 in the equation (4) will depend on the original image [13].

Experiment #1: The characteristic evaluation performances for the sequence of "Lena" images with various types of filtered noise as discussed in Section 2 are illustrated. For this image, we set $f_0=5$. All three images in the sequence have the same MSE value of 100, but their visual qualities are quite different. The results obtained, using the proposed adaptive image quality assessment algorithm, are shown in Table 1.

Table 1. "Lena" Images with Different Filtered Noises

Minimal and Maximal Frequency	MSE	Characteristic evaluation (Q)
$f_{min}=0$, $f_{max}=0.03$	100	0.9120
$f_{min}=0.03$, $f_{max}=0.15$	100	0.6561
$f_{min}=0.5$, $f_{max}=0.707$	100	0.9292

Experiment #2: The characteristic evaluation performances for the sequence of "Lena" images with various types of distortions are illustrated. All images have similar MSE values, but their visual qualities are very different. The performance results are given in Table 2.

Table 2. "Lena" Images with Different Distortions

Type of distortion	MSE	Characteristic evaluation (Q)
Contrast stretching	225	0.9766
Additive white Gaussian noise	225	0.6999
Blurring	224	0.2627
JPEG2000 coding	225	0.1898

Experiment #3: The characteristic evaluation performances for the sequence of "Lena" images distorted by two sources of degradation due to signal dependent and random distortions, respectively, are illustrated. The sources of degradation considered were blurring caused by low-pass filtering and additive white Gaussian noise (zero-mean and standard deviation of σ). The first image contains only blurring, the second image contains some blurring and some additive Gaussian noise and the third image contains only random noise.

All images in the sequence have the same MSE. The characteristic evaluation performance results are given in Table 3.

Table 3. "Lena" Images with Blurring and Additive noise

Image	MSE	Characteristic evaluation (Q)
blurring only	420	0.1395
blurring and additive noise ($\sigma = 17.15$)	420	0.2668
additive white Gaussian noise ($\sigma = 20.55$)	420	0.6273

These wide discrepancies with the subjective quality in the above three experiments illustrate why the MSE or other similar objective measures are not good image quality assessment indicators.

The proposed characteristic evaluation has another interesting property. As can be seen from equation (3), this characteristic evaluation can take both positive and negative values (range of -1 to 1). In this example the reference image is the original "Lena" image and distorted images are inverted original "Lena" image and inverted "Lena" image with additive noise. If the original image is f(m,n), then the corresponding inverted image can be computed as 255-f(m,n). The characteristic evaluation performance results are given in Table 4. Characteristic evaluation values for inverted images have similar magnitudes as the values obtained for the corresponding non-inverted images, but they have a negative sign. This is in a sense related to what we see: images with the same structure but inverted grayscale.

Table 4. "Lena" Images with Inverted Grayscale

Image	MSE	Characteristic evaluation (Q)
Inverted "Lena" image	9258	-0.9955
Inverted "Lena" image with additive noise	9485	-0.6976

Experiment #4: In the last experiment, the algorithm will be applied to the sequence of distorted "Couple" images. All distorted images have same or very close MSE values but their visual qualities are quite different. However, the differences in visual qualities are smaller than in the previous example involving "Lena" images. For this image, parameter f0 is set to 12 cycles/degree because the image has a lot of small objects which attract viewer's attention, which means that information contained in high frequency components is important. If we used $f_0=5$ (as in the case of "Lena" image) for this image, then it would eliminate this information and distortion would be underestimated. On the other hand, if we used $f_0=12$ for "Lena" image, then it would overestimate high frequency distortion in that case. Three different images "Tiffany", "Man" and "Couple" are used to illustrate this point. In these experiments, the proposed image characteristic evaluation produces results which are in good agreement with subjective visual quality of corresponding images. In contrast MSE values are the same for images with very different visual quality.

4. Conclusion

In this paper, a simple image characteristic evaluation algorithm that adapts to various image scenarios by a user-controlled parameter has been presented. First, three main reasons for disagreement between MSE values and subjective visual quality of images have been identified. Then a new quality measure has been introduced. The proposed measure takes into account two HVS properties: nonlinear relationship between intensity and perceived brightness and the presence of spatial filtering in HVS. This measure is based on average value of locally computed correlation coefficients, which is more closely related to the way in which human observer determines quality of an image than a traditional characteristic evaluation such as MSE. Finally, this value is modified by the average value of locally computed correlation coefficients between the original and error images. This way the proposed measure differentiates between random and signal-dependent distortions, which have different effects on human observer. The proposed characteristic evaluation performs reasonably well. The experimental performance results presented here demonstrate that this measure ranks images according to their visual quality in cases when MSE-based measures fail to do that. However, subjective evaluation is still the best way for image quality assessment. HVS is more sophisticated than any mathematically defined characteristic evaluation that we try hard to mimic.

Acknowledgements

References

[1] S. Daly, "The visible difference predictor: an algorithm for assessment of image fidelity in Digital Images and Human Vision", A. B. Watson, MIT Press, Cambridge, MA, **(1993**, pp. 179-206.

[2] R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image dependant quantization and post-quantization data compression", in Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), vol. 3, **(1989)**, pp. 1945-1948.

[3] N. D. Venkata, T. D. Kite, W. Geisler, B. L. Evans and A. C. Bovik, "Image quality assessment based on a degradation model", IEEE, Trans. on Image Processing, vol. 9, no. 4, **(200)** April, pp. 630-650.

[4] T. N. Pappas and R. J. Safranek, "Perceptual criteria for image quality evaluation in Handbook of image and video processing", Academic Press, **(2000)** May.

[5] Z. Wang, A. C. Bovik and L. Lu, "Why is image quality assessment so difficult", in Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), vol. 4, **(2002)**, pp. 3313-3316.

[6] Z. Wang and A. C. Bovik, "Universal image quality index", IEEE, Signal Processing Letters, vol. 9, no. 3, **(2002)** March, pp. 81-84.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment From error visibility to structural similarity", IEEE, Trans on Image Processing, vol. 13, no. 4, **(2004)** April, pp. 600-612.

[8] T. N. Pappas and R. J. Safranek, "Perceptual criteria for image quality evaluation", Handbook of image and video processing, Academic Press, **(2000)** May.

[9] T. N. Pappas, T. A. Michel and R. O. Hinds, "Supra-threshold perceptual image coding", Proc. Int. Conf. on Image Processing (ICIP)", vol. 1, **(1996),** pp. 237-240.

[10] Z. Wang, A. C. Bovik and L. Lu, "Why is image quality assessment so difficult", Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), vol. 4, **(2002)**, pp. 3313-3316.

[11] Z. Wang and A. C. Bovik, "Universal image quality index", IEEE, Signal Processing Letters, vol. 9, no. 3, **(2002)** March, pp. 81-84.

[12] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment From error visibility to structural similarity", IEEE, Trans. on Image Processing, vol. 13, no. 4, **(2004)** April, pp. 600-612.

[13] J. Hwang and J. Cho, "An Simple Processing Algorithm with a User-Controlled Parameter", International Symposium on AACL, **(2013)** November.

Authors

Jeewon Hwang

He received the Ph. D degree from the department of Electronic Engineering, Chonbuk National University in Korea in 1995, He is currently Professor at the Department of Information Technology, Engineering College Chonbuk National University, Jeonju, Korea. His research interests include Digital Signal Processing and Computer Architecture, etc.

Juphil Cho

He received the PhD degree in Electronics Engineering from Chonbuk National University in 2001. From 2000 to 2005, he was a Senior Research Staff at Electronics and Telecommunication Research Institute (ETRI), Daejeon, Korea, where he was involved in the development of the 4G cellular system. He was invited as a Researcher from ETRI in 2006. Since 2005, he has been an associate professor at Department of Radio communication Engineering, Kunsan National University, Gunsan, Korea. His current research interests are the wireless communication technology including spectrum sensing, LTE Advanced, Cognitive Radio, interference detection and cancellation in the wireless local and personal area networks.