A NEW IMAGE DISTORTION MEASURE BASED ON WAVELET DECOMPOSITION

Azeddine Beghdadi

L2TI – Institut Galilée, Univ. Paris 13
93430 Villetaneuse FRANCE
phone: +33 1 45 81 72 18
e-mail : Azeddine.Beghdadi@tsi.enst.fr

Béatrice Pesquet-Popescu

Télécom Paris, Sign. and Im. Proc. Dept.
46, rue Barrault, 75634 Paris, FRANCE
phone: +33 1 45 81 71 92
e-mail : pesquet@tsi.enst.fr

ABSTRACT

This paper introduces a novel image distortion measure based on a non-redundant wavelet decomposition. The proposed image quality measure is compared with the PSNR and a previously introduced Wigner-Ville Distribution-based measure on some real images and simulated degradations. The obtained results are assessed on the basis of consistency with the subjective quality assessment and the complexity of the measure.

1. INTRODUCTION

With the increasing use of applications based on visual data, image quality assessment has become a crucial issue. Since image quality is subjective in nature, its evaluation based on subjective experiments is a widely accepted solution, in addition to the classical PSNR objective measure. However, well established benchmarks for visual quality assessment requires the use of several procedures which have been formalized by ITU recommendations [1]. These procedures are complex, time consuming and not easy to reproduce in a simple research environnement. Furthermore, they cannot be used for online control of image/video quality. It should be also noticed that perfect correlation with the Human Visual System (HVS) could never be achieved due to the natural variations in the subjective quality evaluation. These drawbacks led to the development of other practical and objective measures. Basically, there are two approaches for quantitative image quality measure. The first and more practical one is distortion oriented like the MSE, PSNR and other similar measures. However, for this class of distortion measures, the quality metric is not always correlated with the subjective evaluation for many types of degradations. The second class of evaluation approaches corresponds to HVS-modelling oriented measures. For these two classes the computation of the image quality metric necessitates the reference image. In the last decade, numerous methods for

---

Currently, visiting Professor at ENST, Signal and Processing Departement.

image distortion evaluation inspired from the findings on Human Visual System mechanisms [2] have been proposed. For some known distortions, it is possible to develop a measure which exploits the a priori knowledge on the image degradation. Such approaches are used when some information on the distortion nature and its predictability are well understood. In this case, the computation of the quality metric does not necessitate the reference image. Good results have been obtained for measuring and controlling the distortions generated with block-based compression methods [3]-[4].

This contribution aims at introducing a novel image distortion measure based on wavelet decomposition. Recently, a distortion metric referred to as SNR_W using the joint spatial/spatial-frequency representation has been proved consistent with the subjective evaluation for some distortions [5]. The development of this metric was motivated by the need for a simple distortion metric that is more closely correlated with human visual perception than other simple metrics, such as signal-to-noise ratio (SNR) and root mean-squarred error (RMSE). The Wigner-Ville Distribution (WVD) was used as a part of this metric because it is generally accepted in the vision research community that the early stages of visual processing involve the creation of a joint spatial/spatial-frequency representation [6]. However, the method introduced in [5] is complex and quite time consuming. Furthermore, the 2D-WVD suffers from the disadvantages of aliasing and the presence of interference terms which are unpredictable and non controllable. A windowing is then needed in order to reduce aliasing. To reduce interference terms due to cross terms, a smoothing kernel has to be used. One has to find a trade-off between resolution and interference effect. It is not easy to reach this compromise. It could be also noted that, if the real image is multicomponent, which is usually the case in the applications involving color images, interference terms due to the interaction between any two spectral components of the image (real or complex) at different spatial frequencies makes the spatial/spatial-frequency representation difficult to interpret. Another difficulty is the high dimensionality (4D) of

the 2D-WVD. It is impractical to use full WVD, therefore a Pseudo-WVD is used instead. Furthermore, one has to compute the analytical 2D-signal, which is a difficult task.

All these limitations motivate us to look for another approach, which is based on the two-dimensional Wavelet Transform (WT) [9]. The proposed image quality measure is compared with the PSNR, the Wigner-Ville Distribution-based measure on some real images and simulated degradations. The obtained results are compared on the basis of the subjective quality assessment, the complexity and the relevance for image compression evaluation and control.

The paper is organized as follows: in the next section we briefly introduce our notations. Then, we present the new wavelet-based distortion measure. In Section 4 we compare the results obtained with this evaluation metric with the PSNR and Wigner-Ville Distribution-based measure. We conclude and give some hints for future work in Section 5.

2. NOTATIONS

In this paper, we consider a decomposition on a two-dimensional separable wavelet basis. In the discrete case, the samples of the original discrete image are usually considered to be the approximation coefficients at the finest resolution. The coefficients at rougher resolutions represent then the inner products of the image with functions in the set

$$\bigcup_{j=1}^{j_m} \left\{ \frac{1}{2^j} \psi^d \left(\frac{x}{2^j} - k, \frac{y}{2^j} - l \right), \ (k,l) \in \mathbb{Z}^2 \right\},$$

where $j_m \in \mathbb{N}$ is the highest decomposition level (i.e. coarsest resolution) and d represents the orientation: $d \in \{H, V, D\}$ (horizontal, vertical, diagonal details) if $j \in \{1, \ldots, j_m-1\}$ and $d \in \{H, V, D, A\}$ (A: approximation) if $j = j_m$.

In the sequel, we denote by $\{c_j^d(k,l), \ (k,l)\}$ the wavelet coefficients of such a non-redundant decomposition, at the resolution level j and orientation d.

3. A WAVELET-BASED DISTORTION MEASURE

The criterion we propose, similar to an SNR, for measuring the distortion between an original image f and a distorted version $\widehat{f}$, is defined as:

$$\text{SNR}_{WAV} =$$
$$20 \log_{10} \left(\frac{\sum_{k,l,d} \max_j 2^{-jsp} \left| c_j^d(k_j, l_j) \right|^p}{\sum_{k,l,d} \max_j 2^{-jsp} \left| c_j^d(k_j, l_j) - \widehat{c}_j^d(k_j, l_j) \right|^p} \right)^{1/p}$$
(1)

where $\widehat{c}_j^d$ are the wavelet coefficients of the distorted image $\widehat{f}$ and s is a positive constant. The indices (k_j, l_j) are defined as follows:

$$k_j = \lfloor \frac{k}{2^j} \rfloor, \quad l_j = \lfloor \frac{l}{2^j} \rfloor$$

In other words, the local information around pixel (k,l) in the original image is described by the coefficients (k_j, l_j) for $j \in \{1, \ldots, j_m\}$. Our definition simply indicates that we consider the maximum of the absolute value of the coefficients over the trees spreading all the resolution levels and corresponding to the same spatial location and orientation.

Moreover, one can easily check that the measure $\left(\sum_{k,l,d} \max_j 2^{-jsp} \left| c_j^d(k_j, l_j) \right|^p \right)^{1/p}$ is a norm of the image f. We can be tempted to relate this norm to a Besov norm on the wavelet coefficients, which (provided that the wavelets are regular enough) is equivalent to:

$$\|f\|_{p,q}^s = \left(\sum_{j=1}^{j_m} 2^{-jsq} \left(\sum_{k,l,d} |c_j^d(k,l)|^p \right)^{q/p} \right)^{1/q}$$
$$+ \left(\sum_{k,l} |c_{j_m}^A(k,l)|^p \right)^{1/p}$$
(2)

where p, q, s are the Besov parameters. This norm has been used as a distortion measure [10, 11] to quantify the visual effects of adding noise to an image. The criterion defined by Eq. (1) cannot be reduced to a Besov norm but, if $q = \infty$, (2) becomes

$$\|f\|_{p,\infty}^s = \max_j \left(2^{-jsp} \sum_{k,l,d} |c_j^d(k,l)|^p \right)^{1/p}$$
$$+ \left(\sum_{k,l} |c_{j_m}^A(k,l)|^p \right)^{1/p}.$$

whereas, if $p = \infty$, it reads

$$\|f\|_{\infty,q}^s = \left(\sum_{j=1}^{j_m} 2^{-jsq} \left(\sup_{k,l,d} |c_j^d(k,l)| \right)^q \right)^{1/q}$$
$$+ \sup_{k,l} |c_{j_m}^A(k,l)|.$$

In (1), the use of the maximum of the absolute value of the difference over resolution levels is motivated and inspired by some findings on the nonlinear behaviour of the HVS. Similar nonlinearities have been successfully used to

model intra-cortical inhibition in the primary visual cortex in an HVS-based method for texture discrimination [7].

The weighting factor 2^{-jsp}, depending on the resolution level and on the parameter s, allows us to introduce a rough perceptual masking effect: the influence of low frequency subbands is attenuated in the computation of the quality criterion. We can tune the value of s in order to fit the structure of the image. Moreover, if a stochastic fractal model (2D fBm [12]) of the images is adopted, the p-th order statistical moments of the wavelet coefficients at each resolution level are proportional to $2^{jp(H+1)}$, where $H \in (0,1)$ is the so-called Hurst parameter, which is directly related to the fractal dimension of the image. So, by choosing $s = H + 1$, we can compensate this power-law behaviour. Values of s lower than 1 would correspond to long-range dependent processes.

Note that a criterion similar to (1) could be defined by replacing the wavelet basis by a trame. The potential advantage would be to satisfy the translation invariance property which is often desirable in image analysis tasks. However, this would result in a high increase in computational complexity.

4. SIMULATION RESULTS

We first perform a periodic separable wavelet transform on the maximum decomposition levels: $j_m = \min(\log_2 N, \log_2 M)$, where (N, M) is the image size. We have chosen for simulations only linear-phase wavelets, their symmetry being beneficial for the coefficient localization. The families that have been used are the biorthogonal Daubechies 9/7 and the cubic splines. In all the simulations we have used $p = 2$.

The method has been tested on several types of degradations. Fig. 1 shows the original "parrot" image (256 × 256 pixels) and the corresponding degraded versions when adding Gaussian noise, a grid pattern and when altering by JPEG compression respectively. These images have the same PSNR whereas the human evaluation reveals big differences in quality. A number of observers (more than 25) have been asked to evaluate the subjective quality of these images and rank them according to their own judgement. These results are used as a benchmark for testing the proposed image measure. The results in Table 1 show a strong correlation of the SNR_{WAV} with subjective appreciation of the image quality and also with the previously introduced SNR_W. Note that in these simulations $s = 0.5$, but the same ranking is obtained for all $s \in (0.2, 0.75)$. An additional advantage of criterion SNR_{WAV} over SNR_W is that it can be computed much faster using the classical Mallat algorithm for the computation of the wavelet representation.

5. CONCLUSIONS

We have presented an image quality measure based on the discrete wavelet transform and compared it with a previously introduced Wigner-Ville Distribution measure. For images presenting the same PSNR, the new distortion metric is consistent with subjective quality ranking and less complex than the WVD measure. Further improvements could be achieved if an explicit contrast masking model is incorporated in the design of SNR_{WAV}.

6. REFERENCES

[1] ITU-R Recommendation BT.500-7, "Methodology for the Subjective Assessment of the Quality of Television Pictures", ITU, Geneva, Switzerland, 1995

[2] "Special Issue on Image Quality Assessment", Signal Processing, Vol 70, 1998.

[3] Z. Wang, A.C. Bovik and B.L. Evans, "Blind measurement of blocking artefacts in images", IEEE International Conference on Image Processing, Sep.2000, pp.981-984

[4] L. Meesters and J.B. Martens, "A single-ended blockiness measure for JPEG-coded images", Signal Processing, vol.82, pp.369-38.

[5] R. Iordache, A. Beghdadi, "A Wigner-Ville Distribution-Based Image Dissimilarity Measure", Proc. ISSPA 2001, KL, Malaysia, Aug. 2001, pp.430-433.

[6] T. R. Reed and H. Wechsler, "Segmentation of Textured Images and Gestalt Organisation using Spatial/Spatial-Frequency Relationships", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, p.1-12, 1990.

[7] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms", Journal Opt. Soc. Amer. A, vol.7, pp.923–932, 1990.

[8] S. Daly, "The visual differences predictor : an algorithm for the assessment of image fidelity", in "Digital Images and Human Vision", A.B. Watson, Ed. Cambridge, MA, MIT Press, 1993, pp.179-206

[9] S. G. Mallat, "A Wavelet Tour of Signal Processing", Eds. Academic Press, San Diego, California, 1998

[10] A. Chambolle, R. DeVore, N.Y. Lee, B. Lucier, "Nonlinear Wavelet Image Processing: Variational Problems, Compression and Noise Removal Through Wavelet Shrinkage", IEEE Trans. on Image Processing, vol. 7, no. 3, Mar 1998, pp. 319-335

[11] A. Chambolle, B. Lucier, "Interpreting Translation-Invariant Wavelet Shrinkage as a New Image Smoothing Scale Space", IEEE Trans. on Image Processing, vol. 10, no. 7, July 2001, pp. 993-1000

[12] B. Pesquet-Popescu, J. Lévy Véhel, "Stochastic Fractal Models for Image Processing", IEEE Signal Procesing Magazine, Sept. 2002, pp. 48-62

Fig. 1. Test images. From top to bottom and from left to right: the original image, and its distorted versions by addition of Gaussian noise (PSNR = 23.70), of a grid pattern (PSNR = 23.70) and by JPEG compression (PSNR = 23.74).

Distorsion type/ Quality metric	Gaussian Noise	JPEG compression	Grid pattern
PSNR	23.70	23.70	23.74
SNR_W	21.70	17.66	14.07
SNR_{WAV1}	30.63	29.33	25.75
SNR_{WAV2}	28.98	27.23	24.19
Subjective	I	II	III

Table 1. Comparison of different quality measures on typical distorted images. SNR_W stands for the Wigner-Ville based metric. SNR_{WAV1} corresponds to the criterion in (1) for biorthogonal 9/7 wavelets, while SNR_{WAV2} is the same measure using cubic spline wavelets. Subjective evaluation ranks images from the best quality (I) to the worst one (III).