



**HAL**  
open science

## Wavelet-based euclidean distance for image quality assessment

Stéphane Paris

► **To cite this version:**

Stéphane Paris. Wavelet-based euclidean distance for image quality assessment. International Symposium ELMAR-2010, Sep 2010, Zadar, Croatia. hal-01223691

**HAL Id: hal-01223691**

**<https://hal.science/hal-01223691>**

Submitted on 3 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Wavelet-based Euclidean Distance for Image Quality Assessment

Stéphane Paris

LITA, Univservity, Ile du Saulcy

57000, Metz, FRANCE

Email: paris@univ-metz.fr

**Abstract**—Image quality assessment (QA) mimics the user’s ability in evaluating image distortion. The presented technique uses a wavelet analysis and a distance measure associated to it. As often noticed, the wavelet analysis is a good approximation of the human visual system (HVS). The main contribution of this paper is the definition of a wavelet-based euclidean distance which embeds both, the deep-structure of the images and the specific orientation of every subband. This new definition of distance allows for an effective QA almost parameterless.

## I. INTRODUCTION

This paper deals with full-reference QA ; i.e., the reference image serves to estimate the distortion in the test image. Obviously, the distortion measure must mimic the HVS.

So, based on several key studies, the wavelet analysis is used with the CDF9\_7 filter bank. But, other filters can be used without any modification of the presented technique.

As several researchers already noticed, the wavelet analysis is a good approximation of the HVS. For quantization without perceptual effects, Andrew B. Watson et al. [1] weight the sub-bands of details except the high frequency ones. (i.e., those of the first level of decomposition.) The weights are experimentally estimated. The choice to keep only the sub-bands of details was also based on experimentation where humans were situated in a precisely modeled situation.

Following these results, Emil Dumic et al. [2] shift the weighted wavelet analysis into the QA domain. The analysis is done on the difference between the reference and test images. For every subband, the  $\mathcal{L}_K$ -distance – from the origine (i.e. null distortion) – is measured and weighted.  $K$  is experimentally evaluated. The image quality measure is the sum of all the distances. The main drawbacks of this technique that the authors noticed, concerns the number of parameters and their dependence on the training image database. However, results on LIVE database are promising.

In their SSIM technique [3], Zhou Wang et al. separate the quality measurement into three estimates which evaluate changes in luminance, in chrominance and in structure, with the help of means, standard deviations and correlations of the reference and test images. The product of these three estimates ensures an image quality measure quite stable for different types of distortions : JPEG and JPEG2000 compression, blur, white noise and fast fading. In practice, SSIM is locally estimated for being in adequation with the nonstationary

behavior of the image statistics and of the distortions. Its mean informs on the overall image quality.

Further investigations of the image statistics enabled Hamid Rahim Sheikh et al. [4] to model the HVS and the distortions with random fields (RF). Every subband of the source image is the product of a scalar RF with a Gaussian vector RF. A subband is cut into nonoverlapping blocks of  $M$  coefficients. Every block is used to estimate a unit of the source RF. The distorted image model multiplies the source RF by a scalar RF and adds a white Gaussian noise RF. The image perceived (source or distorted) by the HVS is then approximated by adding a white Gaussian noise RF. As in their previous work, statistics are used to estimate the RF parameters; means, correlations, covariances and eigenvalues. The Visual Information Fidelity criterion (VIF) is the mutual information between the source image and the perceived distorted image relative to the mutual information between the source image and the perceived source image.

The Visual SNR [5] is a two-stage measurement for quantifying the visual fidelity of natural images with the basic tools whom are the wavelet analysis, the root-mean-squared and the threshold function of contrast SNR. The threshold function is parabola shaped. And its parameters are estimated via experiments. The wavelet analysis splits the difference, between the original and distorted images, into subbands whom distortions are measured through the root-mean-squared and the threshold function. The first-stage process determines, with the help of the threshold function, whether the distortion is visible. If it is not, the test image is deemed of be of perfectly visual fidelity and the process stops here. Otherwise, the second-stage process verifies, within a linear combination, the perceived contrast and the global precedence. The global precedence corresponds to distortions observable in the multiscale image’s edges. Thus, the contrast of every subband, relative to the contrast SNR of the images’ difference, is observed.

These are some of the most noticeable works in the field from which the presented work is close ; a CDF9\_7 wavelet analysis simulates the HVS activity. As suggested by the Watson’s study, only the subbands of details are kept, ignoring those of highest frequencies. But the way to arrange the sub-bands coefficients is specific. The quality is measured on every pixel which is described by a coefficient vector provided by the kept subbands.

Note that deeper is the decomposition level, more pixels

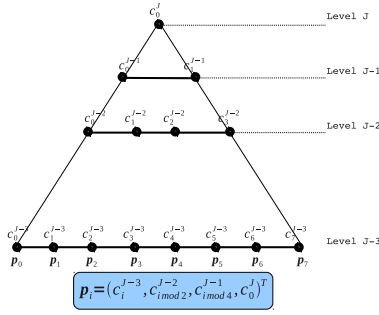


Fig. 1. Implicit localization provided by the wavelet analysis.

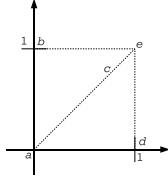


Fig. 2. Illustration of the anisotropic Euclidean distance.

are plausibly concerned by a same coefficient. This is a very simple way to manage a multiscale localization during the quality measurements. Thus, whatever the distance used, a multiscale block grouping is implicitly ensured (see Fig. 1).

Moreover, not only the coefficients of the vectors have to be observed but their orientations too (horizontal, vertical or diagonal). This property is detailed in section II.

From now, every pixel is described by its  $M$ -dimensional *vector descriptor* with  $M = 3 * (J - 1)$  where  $J$  is the deepest decomposition level.

The paper is organized as follow: section II gives the definition and the usefulness of the wavelet-based Euclidean distance; section III explains the experimentations on the LIVE database; section IV concludes and provides plausible future investigations.

## II. WAVELET-BASED EUCLIDEAN DISTANCE

The Euclidean distance between two descriptors  $x = (x_1, \dots, x_n)^T$  and  $y = (y_1, \dots, y_n)^T$ :

$$d_E^2(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$

does not inform about the information delivered by the wavelet analysis. It hypothesizes an isotropic space whereas the wavelet decomposition provides the horizontal, vertical and diagonal details. This fact was already noticed [5].

Thus, it is more pertinent to use an anisotropic Euclidean distance which takes into account the orientation information of the wavelet decomposition. For example, figure 2 shows five 2D points ( $n = 2$ ). The Euclidean distances from point  $a$  to points  $b$ ,  $c$  and  $d$  are equal and the distance to point  $e$  is greater:  $d_E^2(a, b) = d_E^2(a, c) = d_E^2(a, d) = 1$  and  $d_E^2(a, e) = 2$ . In other words, if the information delivered by the two axes have to be considered, the isotropic Euclidean distance is not

appropriate. But, following the study of Wang et al. [6], we can use an anisotropic Euclidean distance:

$$d_{WED}^2(x, y) = \sum_{i,j=1}^n g_{i,j}(x_i - y_i)(x_j - y_j)$$

where the kernel  $g_{i,j}$  is a continuous function positive and decreasing monotonically. For example:  $g_{i,j} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(i-j)^2}{2\sigma^2}}$ . With  $\sigma = 1$  and neglecting the normalization constant, we obtain:  $g_{i,j} \approx e^{-\frac{(i-j)^2}{2}}$ . In this way, the descriptors, which share identical subband coefficients, are close together. With the previous example, the distances are:  $d_{WED}^2(a, b) = d_I^2(a, d) = 1$ ,  $d_I^2(a, c) = 1 + e^{-\frac{1}{2}}$  and  $d_{WED}^2(a, e) = 2 \left(1 + e^{-\frac{1}{2}}\right)$ . If we consider that the abscissa informs on the approximation coefficients and that the ordinate informs on the detail coefficients, the distances  $d_{WED}$  shows that points  $b$  and  $d$  are closer to point  $a$  than is point  $c$ . The distance  $d_{WED}$  emphasizes that the approximation coefficients (resp. details coef.) of  $b$  (resp.  $d$ ) and  $a$  are identical, whereas  $c$  has no common coefficient with the point  $a$ .

With the help of such a distance, the full-reference quality assessment is based on the wavelet analysis of the reference and distorted images. After wavelet analysis, every pixel is described by its descriptor. The mean value of the distances between pixels of the reference and distorted images informs on the global quality.

## III. EXPERIMENTATIONS

Experimentations are made with pictures from Einstein image distortion class and a subset of the LIVE DB: *Caps, Monarch, CoinsInFountain, LightHouse, Parrots, Buildings, Bikes, CarnivalDolls, Plane, Sailing2* and *Statue*<sup>1</sup>. No color information is taken into account and the maximum level of decomposition,  $J$ , is set to 3.

As lighting in table I,  $d_{WED}$  tends to be inversely proportional to SSIM responds.

TABLE I  
EINSTEIN CLASS.

Type	MSE	SSIM	$d_{WED}$
Reference	0	1	0
Mean shifted	144	0.988	0.743432
Contrast	144	0.913	3.93144
Impulse	144	0.840	4.23846
Blurt	144	0.694	5.81087
JPEG	142	0.662	6.70539

Figure 3 shows the results for the *Monarch, Coinsinfountain, Caps* and *Bikes* images. Every subfigure presents the DMOS/WED graphic of all distortion types for a specific image. No outlier appears within these tests; rather, WED distance increases almost linearly with the DMOS increasing.

This is not a surprise since wavelet analysis is known as a good enough method for image texture modeling [1], [5]. By embedding criteria dedicated to other information than

<sup>1</sup>see <http://www.lita.univ-metz.fr/~paris/IQM> for all graphics.



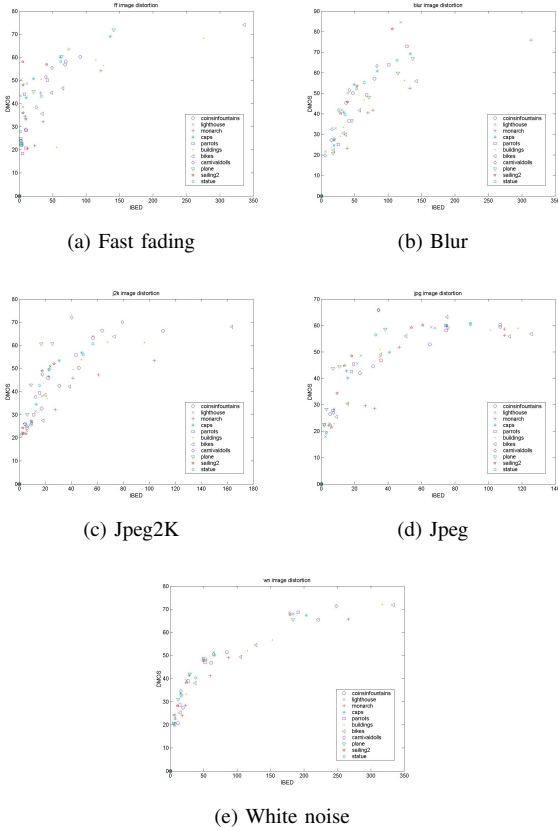


Fig. 5. Every kinds of distortion for all images:  $J = 4$ .

#### IV. CONCLUSION

The present paper proposes an adaptive distance measurement in the context of image quality assessment. The distortion, between the original and test images, is characterized by the subband orientations of a wavelet analysis; every pixel is described, in wavelet domain, by its vector descriptor constituted of the spatial frequencies coefficients which embed a pixel-based neighborhood. With the help of the anisotropic WED distance, both, subband orientations and pixel-based neighborhood are taken into account.

The first experiments encourage us to keep this kind of distance when wavelet analysis is used. More investigations must be done to embed a contrast modification measurement and, maybe, to determine the kind of distortion observed.

#### ACKNOWLEDGMENT

The author would like to thank the LIVE team<sup>2</sup> [8] to make their data set available to the research community free of charge.

#### REFERENCES

[1] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Transaction on Image Processing*, vol. 6, no. 8, pp. 1164–1174, Aug. 1997.

<sup>2</sup><http://live.ece.utexas.edu/research/quality/subjective.html>

[2] E. Dunic, S. Grgic, and M. Grgic, "New image-quality measure based on wavelet," *Journal of Electronic Imaging*, vol. 19, no. 1, 2010.

[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transaction on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[4] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transaction on Image Processing*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[5] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Transaction on Image Processing*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[6] W. L., Z. Y., and F. J., "On the euclidean distance of images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1334–1339, August 2005.

[7] R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Transaction on Information Theory*, vol. 2, no. 38, pp. 713–718, 1992.

[8] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database release 2." [Online]. Available: <http://live.ece.utexas.edu/research/quality>

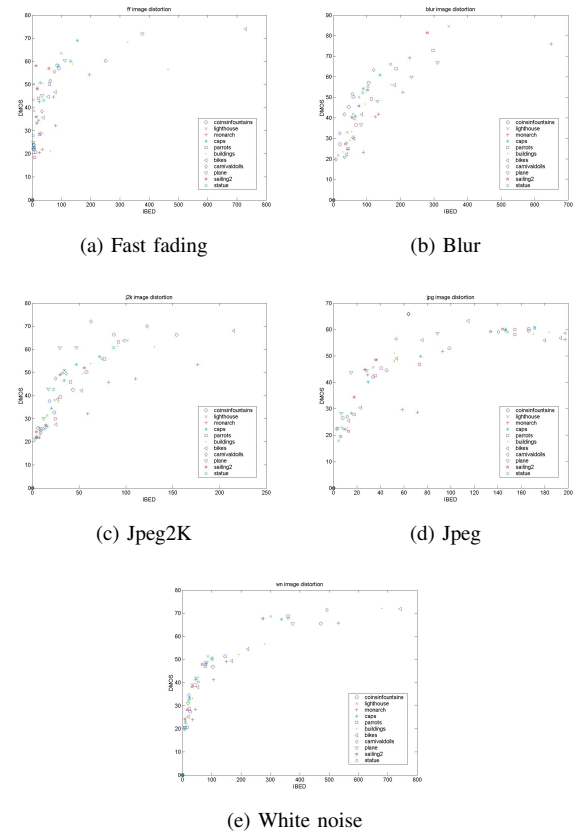


Fig. 6. Every kinds of distortion for all images:  $J = 5$ .

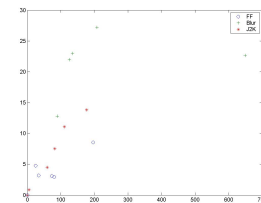


Fig. 7. The best basis selection ( $J = 2 < 8$ ) applied on the fast fading, blur and JPEG2000 distortion of the Monarch image.