



Defining a no-reference image quality assessment by means of the self-affine analysis

Jesús Jaime Moreno Escobar¹ · Oswaldo Morales Matamoros¹ · Ixchel Lina Reyes¹ · Ricardo Tejeida Padilla² · Liliana Chanona Hernández¹

Received: 4 December 2019 / Revised: 16 November 2020 / Accepted: 9 December 2020 /

Published online: 22 January 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

In this paper we propose a novel Blind Image Quality Assessment via Self-Affine Analysis (*BIQSAA*) method by considering the wavelet transform as a linear operation that decomposes a complex signal into elementary blocks at different scales or resolutions. *BIQSAA* decomposes a distorted image into a set of wavelet planes $\omega_{\lambda,\phi}$ of different spatial frequencies λ and spatial orientations ϕ , and it transforms these wavelet planes into one-dimension vector Ω using a Hilbert scanning. From the vector Ω there were obtained their wavelet coefficient fluctuations estimated by the inverse of the Hurst exponent in decibels, whose scaling-law or fractal behavior was obtained by applying Fractal Geometry or Self-Affine Analysis. The scaling exponents calculated for the coefficient fluctuation behavior of Image *Lena* at 24bpp, at 1.375bpp, and at 0.50bpp were $H_{24bpp} = 0.0395$, $H_{1.375bpp} = 0.0551$, and $H_{0.50bpp} = 0.0612$, respectively. Our experiments show that *BIQSAA* algorithm improves in 14.36% the Human Visual System correlation, respect to the four state-of-the-art No-Reference Image Quality Assessments.

Keywords Image quality assessment · Self-affine analysis · Signal processing · Wavelet transform

1 Introduction

The use of digital images has become a natural or usual part of daily-life and it seems to be static and little evolves. The explosion and rapid growth of social networks projected the field of digital image processing at levels never seen.

In 2020, the use of information technology increased dramatically because of the confinement of humans caused by the SARS-COV2 virus pandemic. The change in technological paradigms is almost always due to extraordinary events that changed societies and

✉ Jesús Jaime Moreno Escobar
jomoreno@esimez.mx

¹ Escuela Superior de Ingeniería Mecánica y Eléctrica, Unidad Zacatenco, Instituto Politécnico Naiconal, México, México

² Escuela Superior de Turismo, Instituto Politécnico Naiconal, México, México

human settlements. Thus, the massive use of information media requires the intelligent use of digital resources, transmission media and optimizing the visualization of these images. Therefore, it is important to say that the eye of a human being estimates the quality of a digital image mainly by two factors: i) the analysis of changes in the structure in a static way and ii) the visualization experiences known in the past, which change dynamically over time. At all times that a natural digital image is viewed, an encoded or recovered image is being observed, which is a digital representation of an original image considered without apparent distortion. In most applications, the original image is not showed to human viewer, so estimation of perceptual quality relies solely on the apparent quality of the recovered image. That is, users of digital applications, such as social media, perform a quality evaluation without reference to the original image. Hence, it is important to estimate or predict *a priori* the quality that an image could have, and thus adjust the relevant information that must first be transmitted or stored by digital means.

In this way, the reality is that every day the systems evolve to improve the quality of the expression of a human observer, incorporating these system parameters that correlate better with the human visual system. Like any system, digital image processing requires an input (original image), a processing (image coder), an output (recovered image) and in some cases a feedback to measure how effective the processing step is.

Digital image coders feedback step is known as Image Quality Assessments (IQA) and there are thousand of metrics around the world, Fig. 1. The taxonomy of these quality metrics is divided into three main categories, namely with Reference or R-IQA that needs the original image (Fig. 1a), with reduced reference or RR-IQA that takes some parameters of the input image (Fig. 1b) and No-Reference or NR-IQA that does not take any information from the original source and it is assumed that there is only information about energy, intensity or some other parameter of an image whose quality was presumably distorted (Fig. 1c). In this work we only focus on the NR-IQA field, where there are several algorithms but we only take into account some of them that are considered state-of-the-art assessemnts, such as NIQE [8], ILNIQE [17], ASIQE [4], NRLT [3], and NRSIF [16], described as follows:

- In 2013, Mittal et al. in [8] create the Natural Image Quality Evaluator (NIQE) by measuring deviations from statistical regularities observed in natural images, without knowing previous distortions or human opinions about them., i.e., *completely blind*. NIQE expresses the quality of the distorted image as the gap between the model statistics and those of the distorted image. This NR-IQA method is very useful in training human judgments of known distorted images. This metric is based on the construction of the perceptual quality by mean of the collection of statistical parameters which lies in a Natural Scene Statistic (NSS) model.
- Two years later, Zhang et al. [17] present a variation of NIQE called ILNIQE that integrates natural image statistics parameters derived from multiple process. This algorithm employs a multivariate Gaussian model of some regions of a corpus of pristine natural images.
- In 2017, Gu et al. in [4] apply big data training samples to develop an NR-IQA method called Accelerated Screen Image Quality Evaluator (ASIQE) extracts four types of parameters whose can define complexity of an natural image, such as screen content features, for instance. The main four aspects are image complexity, screen content statistics, global brightness and surface quality, and image sharpness and corners. ASIQE improves the implementation speed, delivering better performance on existing quality assessment databases related to Screen Content Images (SCI).

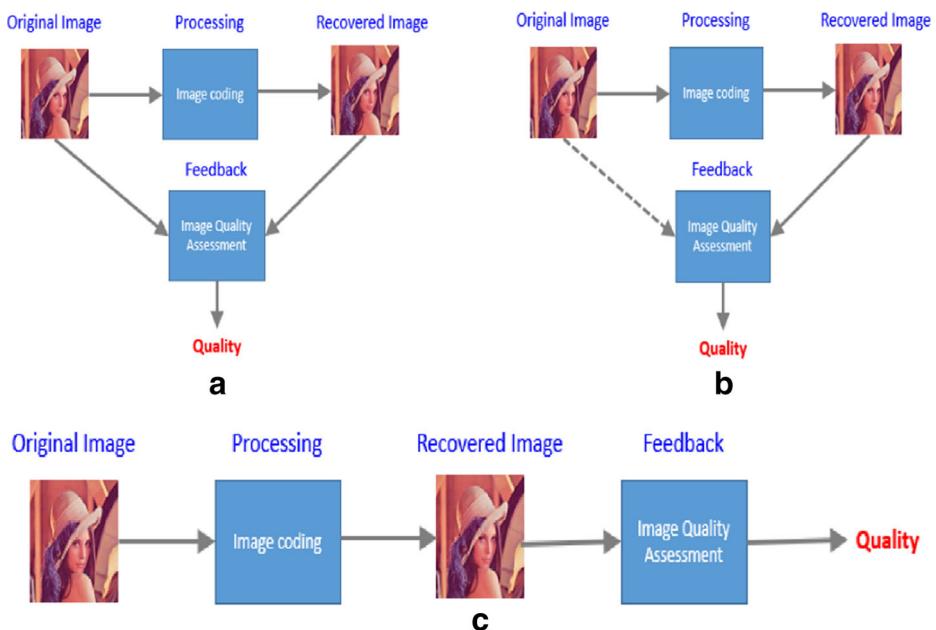


Fig. 1 Image Quality Assessment Taxonomy. **a** with Reference, **b** with Reduced Reference, and **c** No Reference

- In 2018, Fang et al. in [3] generate a NR-IQA metric called No Reference Quality Assessment method by incorporating statistical Luminance and Texture features (NRLT) was based by the perceptual features of the Human Visual System (HVS) because luminance fluctuations and texture properties are parameters those influence in the image understanding. First, they calculate the luminance map through the local normalization in order to extract the statistical luminance features in global scope. Later, they compute gradient maps from the luminance map by using four filters with different orientations, in order to extract the second-order derivatives by local binary pattern. Afterwards, they apply the histogram of high-order derivatives in global scope to extract the texture feature. Finally, they train the mapping function from quality-aware features to subjective ratings via support vector regression, obtaining a better prediction of the visual quality of SCIs.
- Recently, Yang et al. [16] No-Reference image quality assessment via Structural Information Fluctuation (NRSIF) was proposed as a NR-IQA method based on structural information analysis. For each image, their metric calculates the grey-scale fluctuation maps in four detection directions, and thus they estimate the grey-scale fluctuation direction map. Hence, they extract the structural features of each image, in order to collect and transform them into feature vectors. They train their NRSIF model by means of support vector regression for predicting image quality accurately for both natural image and SCI.

There are another latest works in the field of perception where the term image quality varies according to the type of image. For example, Kavitha et al. in [6] used Discrete Wavelet Transform (DWT) and Undecimated DWT (UDTW) to fuse medical images and assessed the quality of the fused image using the metrics such as Spatial Frequency (SF),

Universal Quality Index (UQI), Mutual Information (MI), Information Entropy (IE), Root Mean Square Error (RMSE), and Peak Signal to Noise Ratio (PSNR).

Most of these works make use of artificial intelligence tools, using algorithms that learn how a certain group of images from a certain corpus behaves, i.e., the reported results depend on adequate training of the algorithm. The contribution of Fractal Geometry to the field of Image Quality Assessment, either with or without reference, lies in the fluctuations analysis of the image in frequency domain, i.e., the less fluctuations of pixels (more homogeneity), the less quality of them, avoiding the steps of training in a certain corpus.

Thereby, we start from the hypothesis that the fluctuation of coefficients decreases in the same way that their visual appearance decreases. In addition, this proposal is neither trained with certain parameters nor image collection nor a mathematical model with certain human visual system parameters, in order to speed up its performance. Our algorithm operation is based on measuring the high frequency losses by means of a Self-Affine Analysis. Hence, this work is divided in another three sections. In Section 2, *BIQSAA* algorithm is defined and its results are analyzed in the next section. Finally, in Section *Conclusions*, we express the accomplishment of main objectives.

2 BIQSAA algorithm

From Fig. 2, *BIQSAA* is a NR-IQA algorithm based on the use of an *RGB* input image $\bar{I}(t)$, which is presumably distorted, and is made up of five main steps:

1. Component Transformation of $\bar{I}(t)$ to the YC_bC_r color space obtaining $\bar{I}(t) \rightarrow \hat{I}(t)$ and filtration of the luminance channel of $\hat{I}(t)$ giving as a result $\mathcal{I}_Y(t)$.
2. Transformation of $\mathcal{I}_Y(t)$ to $\mathcal{I}_Y(\omega)$ by means of one level of biorthogonal direct wavelet family, i.e., Cohen–Daubechies–Feauveau wavelet with the low-pass filters of the length 9 and 7 (CDF 9/7).
3. Transformation into an one-dimension and indexed vector Ω by means of a Hilbert Scanning.
4. Filtration the high frequencies from low ones $\Omega \rightarrow \Omega_{\lambda,\phi}$. That is decomposed into a set of wavelet planes $\kappa_{\lambda,\phi}$ of different spatial levels λ (i.e., spatial frequency v) and spatial orientations ϕ , (2).
5. Application of the Fractal Geometry at $\Omega_{\lambda,\phi}$ to estimate the wavelet coefficient fluctuations or Hurst exponent H in decibels.

2.1 Component transformation

The *RGB* color image $\bar{I}(t)$ is decomposed into three color components: Red, Green and Blue color components, since we only need the Luminance information in a gray-scale space. In this work, we use Y , C_b and C_r color components, thus the chrominance channels can be processed independently than luminance one. *BIQSAA* employs an Irreversible Component Transformation (ICT) defined by:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.16875 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

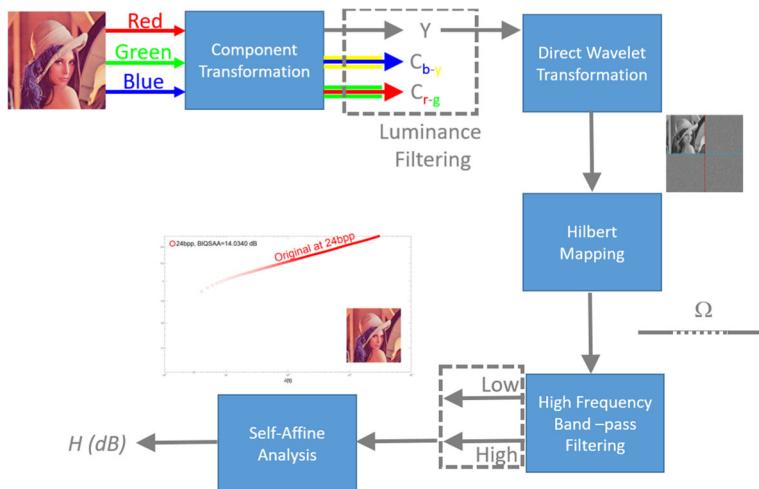


Fig. 2 Main steps of the BIQSAA Algorithm

This gives us as a result the YC_bC_r color space obtaining $\widehat{\mathcal{I}}(t)$ image. A filtration of the luminance channel of $\widehat{\mathcal{I}}(t)$ is performed, giving as a result $\mathcal{I}_Y(t)$.

2.2 Direct wavelet transformation

There are several related works in the field of Artificial Intelligence that make use of various wavelet basis. On the one hand, Kim et al. in [7] develop an algorithm for image segmentation to transform an original feature space into a lower resolution by wavelets to estimate one or multiple feature values as the optimum threshold values. These optimum feature values are projected onto the original feature space to be used to reach the optimum threshold value via a refinement procedure. This metrics implies lower CPU-time used and more accuracy of segmentation. On the other hand, Hong et al. in [5] use the histogram of the oriented gradient and discrete wavelet transform to generate an algorithm for a pedestrian detection. Based on training data via vector machine mechanism, they design a multi-feature for learning, reducing the computational complexity in order to know if a chosen window contains a pedestrian.

The wavelet transform, seen as a mathematical process, is a linear operation that decomposes a complex signal into elementary blocks appearing at different scales or resolutions. These blocks can be generated in a very simple way to simplify their analysis, while allowing to clean the external interference signal. BIQSAA algorithm takes an input image $\widehat{\mathcal{I}}(t)$ as a presumably distorted source and it is decomposed into a set of wavelet planes $\kappa_{\lambda,\phi}$ of different spatial levels λ (i.e., spatial frequency v) and spatial orientations ϕ . It is described as:

$$\mathcal{I}_Y(\omega) = \sum_{\lambda=1}^n \sum_{\phi} \kappa_{\lambda,\phi} + c_n \quad (2)$$

where n is the number of wavelet planes and c_n is the residual plane, and ϕ is the spatial orientation either vertical, horizontal or diagonal. The employed analysis filter is the Daubechies 9-tap/7-tap filter depicted by Table 1.

2.3 Hilbert mapping

In order to generate a time series from a matrix of wavelet coefficients, i.e., transforming a two-dimensional array to one-dimensional array. There are several options, one of them the well-known raster scanning but it does not exploit the redundancy that coefficient neighborhoods have. The fractals, particularly those of Dimension Hausdorff-Besicovitch equal to 2, are always kept in each neighborhood, e.g., the curves of Sierpinski, Moore, Peano, Lebesgue or Hilbert. The latter increases or decreases the row or column in one, never two or more at the same time, resulting in a time series with highly correlated neighboring coefficients, which can be exploited their redundancy.

That is why, both transformed $\kappa_{\lambda,\phi}$ and residual c_n planes are converted into an one-dimension and indexed vector Ω by means of a Hilbert Scanning [9] by exploding the redundancy of the neighborhood. In this way, Algorithm 1 generates a Hilbert mapping matrix θ with level γ , expressing each curve as four consecutive indexes. The level γ of θ is acquired concatenating four different θ transformations in the previous level $\gamma - 1$. Algorithm 1 generates the Hilbert mapping matrix θ , where $\vec{\beta}$ refers a 180 degree rotation of β and β^T is the linear algebraic transpose of β . Thus, each wavelet coefficient at $\mathcal{I}_Y(\omega)$ is stored and ordered, being $\theta_{(i,j)}$ the location index of it into $\Omega_{(1,i \times j)}$.

Algorithm 1 Function to generate Hilbert mapping matrix Ω of size $2^\gamma \times 2^\gamma$

```

Input:  $\gamma$ 
Output:  $\theta$ 
1 if  $\gamma = 1$  then
2    $\theta = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ 
3 else
4    $\beta = \text{Algorithm 1 } (\gamma - 1)$ 
5    $\theta = \begin{bmatrix} \beta^T & (\vec{\beta})^T + (3 \times 4^{\gamma-1}) \\ \beta + 4^{\gamma-1} & \beta + (2 \times 4^{\gamma-1}) \end{bmatrix}$ 

```

The initial level of the θ matrix (Algorithm 1, Line 2) shows an *Up* type movement, that is, start at the top and end at the top [1 → 2 → 3 → 4]. Therefore, the scan of the wavelet coefficients starts in the residual plane and ends by horizontal details.

Table 1 9/7 Analysis Filter

Analysis Filter

i	Low-Pass Filter $h_L(i)$	High-Pass Filter $h_H(i)$
0	0.6029490182363579	1.115087052456994
±1	0.2668641184428723	-0.5912717631142470
±2	-0.07822326652898785	-0.05754352622849957
±3	-0.01686411844287495	0.09127176311424948
±4	0.02674875741080976	

2.4 High frequency band-pass filtering

The decomposition of the $\mathcal{I}_Y(t)$ signal via a Wavelet transform is described in a quadtree or $\mathcal{I}_Y(\omega)$, i.e., there are low-frequency planes in the residual one c_n which is still in pixel domain while in frequency domain, high-frequency planes in horizontal, vertical, or diagonal spatial orientations are presented. The human eye is more sensitive to low frequency loss since it finds the information of the image shape, for example. Thus, image coders seek to eliminate irrelevant details to be perceived by the human eye; these eliminated details are found in large and sometimes redundant areas of high frequencies. That is why the apparent poor quality of an $\mathcal{I}_Y(t)$ signal is due to loss of high-frequency planes.

Therefore, we denote $2^y \times 2^y$ and 4^y the resolution of matrix $\mathcal{I}_Y(\omega)$ and after being indexed vector $\Omega(\omega)$, respectively. With the aim of filtering the high frequencies from low frequencies, we eliminate the first 4^{y-n} coefficients, namely, the residual plane c_n at the n level, Fig. 3. This procedure generates the vector $\Omega \rightarrow \Omega_{\lambda,\phi}$, which only contains high frequencies with dimension equal to $3 \times 4^{y-n}$, i.e., the present algorithm estimates the no-reference quality in the areas where presumably there were visually important losses of information, using only 75% of the high-frequency information.

2.5 Self-affine analysis

In many real-world complex systems emerge patterns resulting from the repeated non-linear interactions among its constituents over time intervals, which are short compared to the characteristic time scales of their prior evolution. Fractal geometry is a mathematical tool created to characterize, model and predict the complex systems' behavior with scale-invariant or self-similarity symmetry usually characterized by power-laws with non-integer scaling exponents.

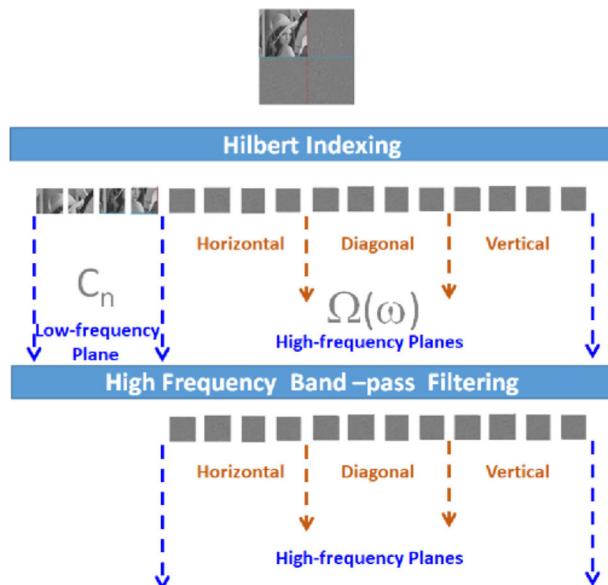


Fig. 3 High frequency band-pass filtering

Self-similarity is the basic feature fractals: a dilated part of a fractal by the same magnification factor in all directions cannot be distinguished from the original. However, random fractals are generated using a stochastic process. The structure of a random or self-affine fractal is invariant under the anisotropic magnification.

In this way, the main goal of this proposal is to apply the Fractal Geometry to estimate the wavelet coefficient fluctuations, in order to develop a novel No-Reference Image Quality Assessment.

A self-affine time series is scale-invariant under the anisotropic transformation $x \rightarrow bx$, $y \rightarrow ay$. Therefore, propose to define that self-affine time series of the wavelet coefficient fluctuations by: we

$$\Omega_{\lambda,\phi}(bx) = a\Omega_{\lambda,\phi}(x) \equiv b^H \Omega_{\lambda,\phi}(x), \quad (3)$$

where the scaling exponent $H = \frac{\log a}{\log b}$, is called the Hurst exponent [1].

The solution of the functional (3) is the scaling law $\Omega_{\lambda,\phi}(x) = Ax^H$. This power law describes the behavior of the fluctuation $\Omega_{\lambda,\phi}$ as function of the frequency scale parameter s at least asymptotically: $\Omega_{\lambda,\phi}(s) \sim s^H$. The power law should be valid for at least one order of magnitude of s values.

Wavelet coefficients $\Omega_{\lambda,\phi}$ generated by complex systems display fluctuations fitted by scaling laws over a wide range of frequency scales and with one fractal scaling exponent, showing the state of the system in one dimension. These fluctuations can be studied quantitatively such as self-affine time series. In this way, the Hurst exponent points out whether the time series of fluctuations display correlations or not.

When $H > 1/2$, self-affine time series are persistent (positive correlations in fluctuations), i.e., the last value of a rising trend is more probable to be followed by a bigger value, and vice versa. When $H < 1/2$, self-affine time series are antipersistent (negative correlations in fluctuations), i.e., the last value of a rising trend is more probable to be followed by a smaller value, and vice versa. Lastly, when $H = 1/2$, self-affine time series are random walk (no correlations in the fluctuations).

The persistent/antipersistent behavior of high frequency wavelet coefficients holds for all frequency scales, where the self-affinity relation [1] holds, until a crossover s_x appears. A crossover is a specific time where the scaling law behavior changes, i.e., there is a scaling exponent for small scale parameters and another one for large scale parameters. This particular time is denoted by its characteristic scale parameter.

High-frequency correlated fluctuations Δx_i are characterized by a finite characteristic correlation decay frequency ω_x , leading to a crossover in the scaling behavior of the integrated series $x_i = \sum_{j=1}^i \Delta x_i$. The position of the crossover over time might be numerically different from ω_x , thus we denote it by s_x . While $H > 1/2$ is observed on low frequency scales (persistent behavior), the asymptotic behavior for high-frequency scales $s \gg \omega_x$ and $>> s_x$) is always characterized by $H = 1/2$ (antipersistent behavior) or $H < 1/2$ (random walk behavior), since all correlations have decayed. High-frequency correlations followed by scaling low-frequency correlations characterize many natural time series, e.g., for particular general weather situations in temperature data and due to respirational effects in heartbeat data, there have been found short-term correlations.

Crossovers s_x^n in the scaling behavior of $\Omega_{\lambda,\phi}$ can be yielded by different regulation mechanisms on fast and slow pixel transformed scales $\omega_{\lambda,\phi}$. Fluctuations of loss of quality display different scaling behavior on frequency scales. Figure 4 shows the coefficient fluctuation behavior of Image *Lena* at 24bpp in red, at 1.375bpp in green, and at 0.50bpp in blue, using JPEG Distortion. Thus, we estimated $H_{24\text{bpp}} = 0.0395$, $H_{1.375\text{bpp}} = 0.0551$, and $H_{0.50\text{bpp}} = 0.0612$ using (3). The lower H , the better quality. Hence, (4), H is inverted

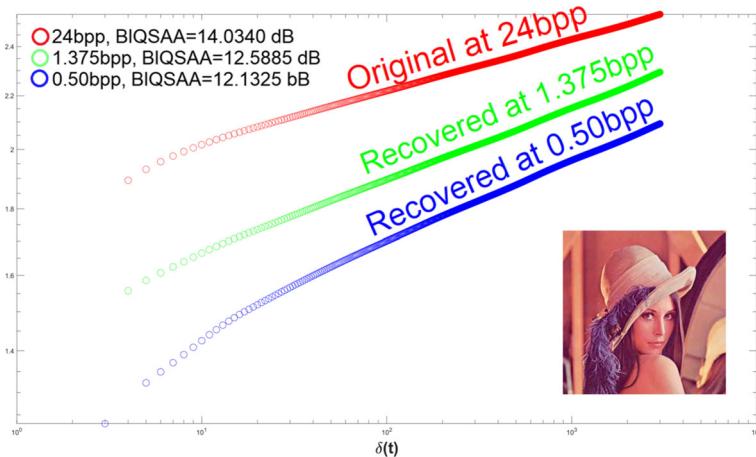


Fig. 4 Coefficient fluctuation behaviour of the Image *Lena* at 24bpp in red, at 1.375bpp in green, and at 0.50bpp in blue, JPEG Distortion

and expressed in decibels (dB). In this way, *BIQSAA* also is defined with this equation and it is established as an Image Quality Assessment that does not need any reference image, since for estimating the fluctuation behavior only the coefficient fluctuation is employed.

$$BIQSAA = 10 \log_{10} (H)^{-1} dB \quad (4)$$

3 Experimental results and analysis

This proposal is compared with a corpus which contains several hundred of original and distorted from the following five Image Databases:

1. LIVE [11]: Image Database of the Laboratory for Image and Video Engineering (LIVE). This corpus contains the subjective evaluation of 161 people to 982 images (29 source images) which were distorted by Noises: JPEG2000, JPEG, White noise, Gaussian blur, and Fast-fading wireless recovered this study generated 25000 individual evaluations. Figure 5 shows a sample of corpus presented to human observers. It was developed by the University of Texas at Austin, USA.

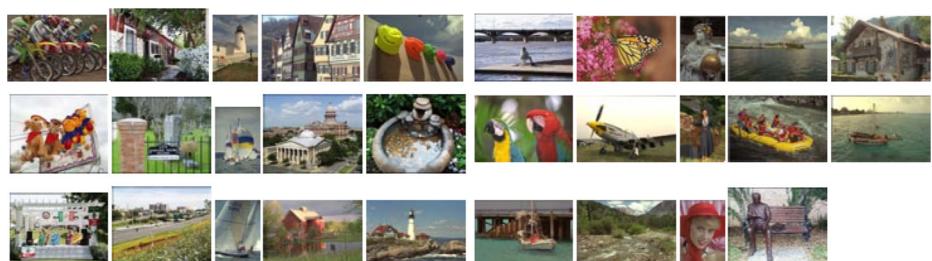


Fig. 5 Image Database of the Laboratory for Image and Video Engineering (LIVE): All 29 Source images [11]



Fig. 6 Categorical Subjective Image Quality (CSIQ) Image database: All 30 Source images [2]

2. CSIQ [2]: Categorical Subjective Image Quality (CSIQ) Image database was jointly developed by the University of Washington and Oklahoma State University, USA. CSIQ has 30 source images (Fig. 6), obtaining 5000 evaluations from 25 observers. The source images were distorted at 6 levels with the following noises: Gaussian blurring, additive white Gaussian, additive pink Gaussian, global contrast decrements, JPEG2000, and JPEG, resulting in this image database 866 distorted images to be evaluated.
3. TID2013 [10]: Tampere Image Database version 2013 (TID2013) was developed by the Tampere University of Technology in Finland. TID2013 contains 25 original images, and 5 out of the 24 types of noise were applied, resulting in up to 3000 distorted images to be evaluated. To find the Mean Opinion Scores (MOS), up to 985 subjective evaluations were processed and obtained, Fig. 7.
4. QACS [14]: In 2016, Quality Assessment of Compressed Screen Content Images (QACS) Database was proposed at Object Search Laboratory of Nanyang Technological University, Singapore. This Image Database considers 24 original images, Fig. 8. QACS compensates some characteristics missing in SIQAD. Authors add High Efficiency Video Coding distortion along with the most common distortions such as layer segmentation-based coding, JPEG2000 and JPEG coders, motion blur, contrast change, Gaussian noise, and Gaussian blur.
5. SIQAD [15]: In 2015, Screen Content Image Quality Assessment Database (SIQAD) was developed by Nanyang Technological University, Singapore. This image database has 980 distorted Screen Content Images (SCI) from 20 pristine SCIs. This database presents a methodology of 11 scales to obtain three perceptual evaluations such as: pictorial regions, textual, and observing the entire image. The distortions or noise



Fig. 7 Tampere Image Database (TID2013): All 25 Source images [10]

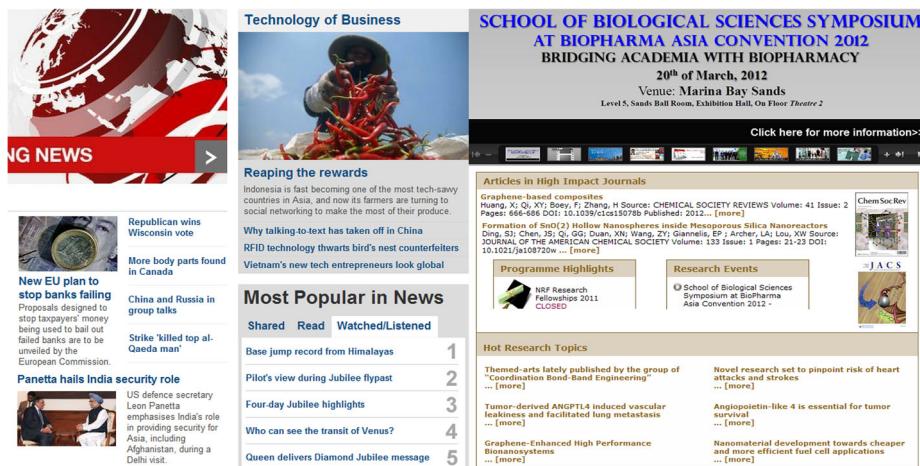


Fig. 8 Quality Assessment of Compressed Screen Content Images (QACS) Database: All 24 Source images [14]

presented are: Artifacts due to a compression coder, Gaussian blurring and Contrast Change. Figure 9 shows some examples of source images from this image database.

In addition, the performance of the proposed algorithm is compared with five No-Reference Image Quality Assessment (NR-IQA) algorithms considered as the most relevant metrics in the field:

1. NIQE [8],
2. ILNIQE [17],
3. ASIQE [4],
4. NRLT [3], and
5. NRSIF [16].

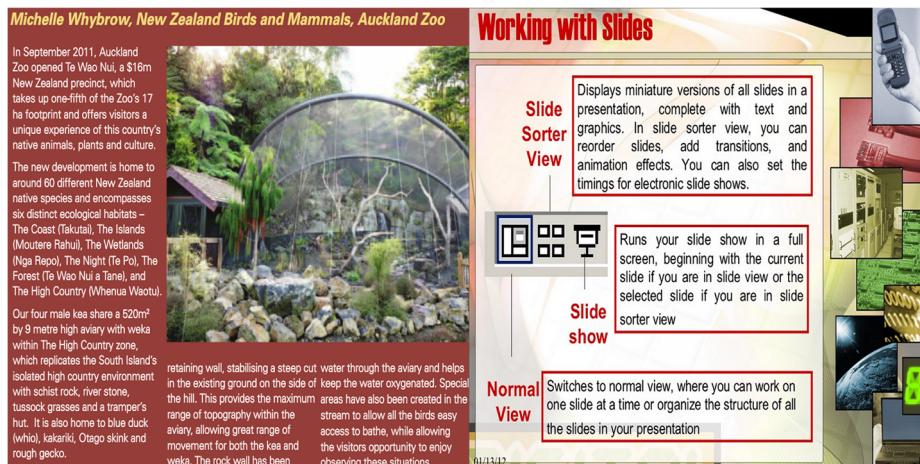


Fig. 9 Screen Content Image Quality Assessment Database (SIQAD): Sample of Source images [15]

Main features of these NR-IQA were discussed in *Section 1. Introduction*.

Image databases contain many images distorted by different noises at different levels. Each of these images are evaluated by many human observers, until getting many evaluations per image. The average of the results of these evaluations are known as the Most Opinion Score (*MOS*). Thus, each image has an associated *MOS* considered as the ground truth of the perception of the image and a MOS_i as the score set of all images of the database. In the same way, any NR-IQA evaluates each image and obtains its own evaluation known as the Predicted Opinion Score (*POS*) and the assessment set of all images of the database known as POS_i . All five Image Databases contain the evaluation results of a observer group or MOS_i and the one of NR-IQA or POS_i are, both are normalized to the scope from 0 to 1, according to following equation:

$$\hat{v}_i = \frac{v_i - \min(v_i)}{\max(v_i) - \min(v_i)} \quad (5)$$

where v can be changed by MOS_i or POS_i , depending on what evaluation needs to be normalized, which are founded after predicting the image quality across all five Image Databases.

The Strength of Relationship (SR) indicates how related are two effects to trend or not to the same response. In this experiment, we use the following Performance Measures (PM):

- Spearman Rank Ordered Correlation Coefficient (SROCC),
- Linear Correlation Coefficient (LCC),
- Kendall Rank Ordered Correlation Coefficient (KROCC),
- Root-Mean-Squared Error (RMSE, see (6)), and
- Mean Absolute Error (MAE, see (7)).

$$RMSE = \sqrt{\sum_i^N \frac{(Predicted_i - Actual_i)^2}{N}} \quad (6)$$

$$MAE = \sum_i^N \left| \frac{Predicted_i - Actual_i}{N} \right| \quad (7)$$

For any correlation coefficient, any SR value which tends to 1 means a high correlation with human visual system. Furthermore, the lower RMSE and MAE , i.e. that tends to zero, the better, this reflects the best precision of a given NR-IQA. So, we compare SR and a normalized MOS_i or POS_i giving as a result a Performance Measure.

On the one hand, both SROCC and LCC are one of the correlation coefficients to measure SR, widely used in the literature, not only in the Image Quality Assessment but also in the field of Computer Vision. The fast processing of LCC gives us a very accurate estimate when the populations are from the same space, e.g., PSNR vs SNR (in decibels) or MAE vs MSE (mean error). When the populations are not in the same space, a non-parametric coefficient such as SROCC is used, because it does not reflect the SR of the compared variables. Both correlation coefficients have problems when the populations do not have the same space. The clearest example is PSNR vs MSE, both estimators are the same metric in essence, but in different space (decibels vs error) both LCC and SROCC estimate overestimated correlations of the same result. These impressions are corrected by KROCC, reflecting the strength of the relationships of the compared variables. Therefore, the metrics with the same essence but different space obtain the same correlation coefficient, however, the main problem is the processing time due to its complexity. We present

the three-correlation coefficients so that the present metric can be compared with other estimators not considered in this study.

On the other hand, Table 2 shows all assessment results of the Strength of Relationship and Error of all Performance Metrics across the five Image Databases. Also, Table 2 shows the comparison the average performance of this proposal against the latest algorithms in the field of No-Reference Image Quality Assessment, highlighting the following:

- ***BIQSAA=0.8688*** is the best metric in terms of SROCC followed by NRSIF=0.8429, NRLT=0.7622, ASIQE=0.6366, ILNIQE=0.5855, and NIQE=0.5362. Thus, this proposal improves by 29.15% the average of all other metrics.

Table 2 Performance comparisons on LIVE [11], CSIQ [2], TID2013 [10], SIQAD [15], and QACS [14] Image Databases

Image Database	Performance Measure	NIQE [16]	ILNIQE [16]	ASIQE [16]	NRLT [16]	NRSIF [16]	<i>BIQSAA</i> PROPOSAL
LIVE [11]	SROCC	0.8459	0.8571	0.6936	0.9060	0.9417	0.9597
	LCC	0.8542	0.8648	0.7057	0.9101	0.9441	0.9621
	KROCC	0.6906	0.6999	0.5480	0.7777	0.7919	0.8575
	RMSE	8.6874	8.4125	11.6255	7.0391	5.2488	6.9428
	MAE	6.6564	6.6617	9.2627	5.3025	3.8963	3.8395
CSIQ [2]	SROCC	0.6367	0.8241	0.5116	0.6615	0.8642	0.8822
	LCC	0.7268	0.8635	0.6477	0.6837	0.8857	0.9037
	KROCC	0.5194	0.6807	0.4016	0.5367	0.6820	0.7497
	RMSE	0.2419	0.1956	0.2611	0.2528	0.1240	0.1215
	MAE	0.2079	0.1692	0.2288	0.2149	0.0981	0.0942
TID2013 [10]	SROCC	0.3708	0.5527	0.4024	0.6300	0.7363	0.7543
	LCC	0.4580	0.4580	0.6473	0.6190	0.7856	0.8036
	KROCC	0.2708	0.4080	0.3269	0.4721	0.5538	0.5058
	RMSE	1.1956	1.0613	1.0859	0.9914	0.7711	0.9062
	MAE	1.0005	0.8701	0.9219	0.8143	0.5849	0.8078
QACS [14]	SROCC	0.4002	0.3152	0.7594	0.7637	0.8618	0.8798
	LCC	0.4787	0.3449	0.7635	0.8024	0.8770	0.8950
	KROCC	0.3172	0.2344	0.5771	0.6060	0.6892	0.6295
	RMSE	2.0724	2.1848	1.6332	1.5303	1.0188	1.4340
	MAE	1.7728	1.8688	1.2870	1.2258	0.8076	1.6000
SIQAD [15]	SROCC	0.4275	0.3786	0.8159	0.8500	0.8107	0.8680
	LCC	0.4421	0.4455	0.8473	0.8793	0.8375	0.8973
	KROCC	0.3129	0.3060	0.6198	0.6624	0.6225	0.7509
	RMSE	13.2802	13.2601	8.8653	8.1537	7.7030	8.1081
	MAE	10.9056	10.8375	7.1224	6.4788	6.2362	6.9094
Average	SROCC	0.5362	0.5855	0.6366	0.7622	0.8429	0.8688
	LCC	0.5920	0.5953	0.7223	0.7789	0.8660	0.8923
	KROCC	0.4222	0.4658	0.4947	0.6110	0.6679	0.6987
	RMSE	5.0955	5.0229	4.6942	3.5935	2.9731	3.5025
	MAE	4.1086	4.0815	3.7646	2.8073	2.3246	2.6502

Bolds indicate the best results, while italics the second best

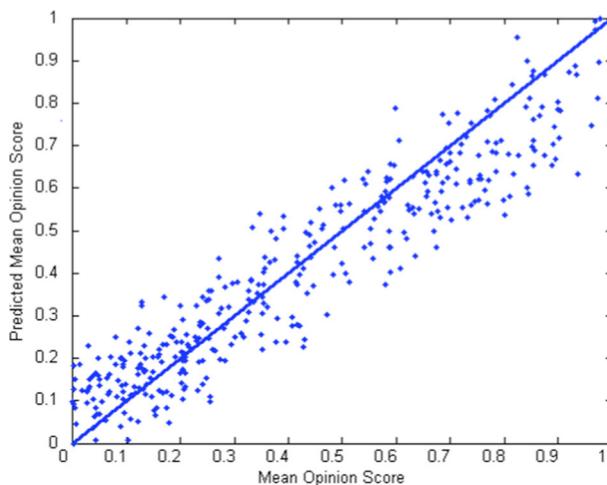


Fig. 10 MOS_i vs POS_i (both normalized). POS_i is predicted by BIQSAA in LIVE [11] image data base

- In the same way and order for LCC, our proposal gets **0.8923**, while the other NR-IQAs get the following: NRSIF=0.8660, NRLT=0.7789, ASIQUE=0.7223, ILNIQE=0.5953, and NIQE=0.5920. So, this proposal improves on the average from 3.04% to 50.74% of the rest of assessments.
- **KROCC=0.6987** while NIQE=0.4222, ILNIQE=0.4658, ASIQUE=0.4947, NRLT=0.6110 and *NRSIF*= 0.6679. It means that our proposal increases SR in 4.61% regarding NRSIF, which is the second best NR-IQA.
- Descending order of the results, in terms of RMSE from the worst to the best, is as follows: NIQE=5.0955, ILNIQE=5.0229, ASIQUE=4.6942, NRLT=3.5935, **BIQSAA=3.5025**, and NRSIF=2.9731. These results leads us to note that this proposal is the second best metric.
- The same order is found when MAE was used as a performance metric, as follows: NIQE=4.1086, ILNIQE=4.0815, ASIQUE=3.7646, NRLT=2.8073, **BIQSAA=2.6502**, and NRSIF=2.3246.

We use a scatter plot for depicting the relationship between subjective results (normalized MOS_i) and objective results (normalized POS_i) in the particular case of the LIVE Image Database. According to Fig. 10, the scatter points of the individual are highly concentrated around the perfect linear correlation (diagonal line).

4 Conclusions

In the present manuscript we present a No-Reference Image Quality Assessment (NR-IQA) via Self-Affine Analysis called **BIQSAA**, which improves existing NR-IQAs in the field, regarding correlation with the human visual system in 0.0258 (SROCC), 0.0263 (LCC) and **0.0308** as for the average Strength of Relationship. Regarding the estimation of mean error, either RMSE or MAE, this proposal is the second best NR-IQA. Despite of being the best No-Reference metric, it is important to note that the present proposal only reaches 69.87% correlation with the human visual system and it has a low performance because observers,

employed in the experimental assessment, make evaluations comparing the original images against the distorted ones, which would be corrected if there was a database especially for NR-IQAs. The effectiveness of this proposal lies in the employment of a fractal or Self-Affine Analysis in the Image Quality Assessment prediction. The future of this research is to propose a Reference Image Quality Assessment for both two and three dimensional. It is important to mention, this work considers as basis function the Cohen-Daubechies-Feauveau wavelet, since it is used in compressors to allow progressive transmission (e.g., JPEG200), and therefore, it can not only propose metrics to assess Image Quality without reference, but also metrics to be completely embedded and to estimate the quality with the bits received progressively. In addition, the introduction of another fractal analysis such as multivariate one. If a complex system exhibits scaling behavior, it could be suggested what kind of scaling universality class it belongs to. The universality hypothesis makes able to propagate the predictive power from a simple model to more complex situations and entire classes of stochastic processes, since the system can often be treated as being composed of many independent sub-systems.

The future of this research will include evaluation parameters already known as the Universal Quality Index (UQI) and Detrended Fractal Analysis (DFA), in order to weigh the findings found in this work. Hence, this proposal may evolve to be tested in different types of sources such as medical images and even spectral ones, according to Kavitha et al. in [6].

Furthermore, in this work it has mentioned five databases for assessing quality of natural images, such as LIVE, CSIQ, TID2013, QACS, and SIQAD. These Image Databases provide original images and recovered ones, distorted by a given noise. For human observers, this will not make much difference because they are near-duplicate images. There are another kind of Image Databases which have near-duplicate images created by changing either intensity values or very small portions of the original images, this effect may not be distinguished easily by human observers and Most Opinion Score provided by human observers could fail. Specifically, in the medical imaging field these differences could make an important difference in the estimated Most Opinion Score. Then, the proposal can be tested in several image databases that meet these requirements, Thyagarajan in [13] analyzes several of these databases. With this, the number of Image Quality Assessment Algorithms should be expanded, such as the one proposed by Thyagarajan in [12], where many of these databases were used.

Acknowledgment This article is supported by Instituto Politécnico Nacional of Mexico by means of projects No. 20200638, 20200324, and 20202061 and granted by Secretaría de Investigación y Posgrado, and Consejo Nacional de Ciencia y Tecnología of Mexico. The research described in this work was carried out at Escuela Superior de Ingeniería Mecánica y Eléctrica of the Instituto Politécnico Nacional, Campus Zacatenco. The authors of this article thank Lic. Pedro Arrechea Alfaro (Head of the Human Resources Department) the for his support in carrying it out.

Conflict of interests The authors declare that they have no conflict of interest.

References

1. BB M (1999) Multifractals and 1/f noise: wild self-affinity in physics. Springer, Berlin
2. C. Larson E, Chandler D (2010) Most apparent distortion: Full-reference image quality assessment and the role of strategy. J. Electronic Imaging 19:011006. <https://doi.org/10.1117/1.3267105>

3. Fang Y, Yan J, Li L, Wu J, Lin W (2018) No reference quality assessment for screen content images with both local and global feature representation. *IEEE Trans Image Process* 27(4):1600–1610. <https://doi.org/10.1109/TIP.2017.2781307>
4. Gu K, Zhou J, Qiao J, Zhai G, Lin W, Bovik AC (2017) No-reference quality assessment of screen content pictures. *IEEE Trans Image Process* 26(8):4005–4018. <https://doi.org/10.1109/TIP.2017.2711279>
5. Hong G-S, Kim B-G, Hwang Y-S, Kwon K-K (2016) Fast multi-feature pedestrian detection algorithm based on histogram of oriented gradient using discrete wavelet transform. *Multimedia Tools and Applications* 75(23):15229–15245. <https://doi.org/10.1007/s11042-015-2455-2>. <http://link.springer.com/10.1007/s11042-015-2455-2>
6. Kavitha S, Thyagarajan KK (2017) Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation. *Soft Comput* 21(12):3307–3316. <https://doi.org/10.1007/s00500-015-2009-6>. <http://link.springer.com/10.1007/s00500-015-2009-6>
7. Kim B-G, Shim J-I, Park D-J (2003) Fast image segmentation based on multi-resolution analysis and wavelets. Elsevier Science Inc., Amsterdam. [https://doi.org/10.1016/S0167-8655\(03\)00160-0](https://doi.org/10.1016/S0167-8655(03)00160-0)
8. Mittal A, Soundararajan R, Bovik AC (2013) Making a “completely blind”? image quality analyzer. *IEEE Signal Processing Letters* 20(3):209–212. <https://doi.org/10.1109/LSP.2012.2227726>
9. Moreno J, Morales O, Tejeida R, Posadas J, Quintana H, Sidorov G (2019) Distributed learning fractal algorithm for optimizing a centralized control topology of wireless sensor network based on the hilbert curve l-system. *Sensors* 19(6):1442. <https://doi.org/10.3390/s19061442>. <https://www.mdpi.com/1424-8220/19/6/1442>
10. Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F, Kuo C-CJ (2015) Image database tid2013: Peculiarities, results and perspectives. *Signal Process Image Commun* 30:57–77. <https://doi.org/10.1016/j.image.2014.10.009>
11. Sheikh HR, Sabir MF, Bovik AC (2006) A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans Image Process* 15(11):3440–3451. <https://doi.org/10.1109/TIP.2006.881959>
12. Thyagarajan KK, Kalaiarasi G (2018) Pulse coupled neural network based near-duplicate detection of images (PCNN - NDD). *Advances in Electrical and Computer Engineering* 18(3):87–96
13. Thyagarajan KK, Kalaiarasi G (2020) A review on near-duplicate detection of images using computer vision techniques. *Archives of Computational Methods in Engineering*. <https://doi.org/10.1007/s11831-020-09400-w>. <http://link.springer.com/10.1007/s11831-020-09400-w>
14. Wang S, Gu K, Zhang X, Lin W, Zhang L, Ma S, Gao W (2016) Subjective and objective quality assessment of compressed screen content images. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 6(4):532–543. <https://doi.org/10.1109/JETCAS.2016.2598756>
15. Yang H, Fang Y, Lin W (2015) Perceptual quality assessment of screen content images. *IEEE Trans Image Process* 24(11):4408–4421. <https://doi.org/10.1109/TIP.2015.2465145>
16. Yang X, Wang T, Ji G (2020) No-reference image quality assessment via structural information fluctuation. *IET Image Process* 14(2):384–396
17. Zhang L, Zhang L, Bovik AC (2015) A feature-enriched completely blind image quality evaluator. *IEEE Trans Image Process* 24(8):2579–2591. <https://doi.org/10.1109/TIP.2015.2426416>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.