



HAL
open science

No-Reference Learning-based and Human Visual-based Image Quality Assessment Metric

Christophe Charrier, Abdelhakim Saadane, Christine Fernandez-Maloigne

► **To cite this version:**

Christophe Charrier, Abdelhakim Saadane, Christine Fernandez-Maloigne. No-Reference Learning-based and Human Visual-based Image Quality Assessment Metric. 19th International Conference on Image Analysis and Processing, Sep 2017, Catania, Italy. hal-01595943

HAL Id: hal-01595943

<https://hal.science/hal-01595943>

Submitted on 27 Sep 2017

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.

No-Reference Learning-based and Human Visual-based Image Quality Assessment Metric

Christophe Charrier¹, Abdelhakim Saadane², and Christine Fernandez-Maloigne³

¹ Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

² Université de Nantes, XLIM Lab

³ Université de Poitiers, XLIM Lab

Abstract. With the rapid growth of multimedia applications and technologies, objective image quality assessment (IQA) became a topic of fundamental interest. No-Reference (NR) IQA algorithms are more suitable to real-world applications where the original image is not available. In order to be more consistent with human perception, this paper proposes a new NR-IQA metric where the input image is firstly decomposed to several frequency sub-bands which mimic the human visual system (HVS). Then, the statistical features are extracted from these frequency bands and used to fit a multivariate Gaussian distribution (MVG). Finally, the model obtained by training predicts the quality of the input image. Experimental results demonstrate the method effectiveness and show its robustness when tested by different databases. Moreover, the predicted quality is more consistent with human perception.

Keywords: image quality assessment, no reference, human visual system, visual sub-bands, multivariate Gaussian distribution

1 Introduction

From the last two decades, a great number of researches have been conducted to design robust No-Reference Image Quality Assessment (NR-IQA) algorithms. As they don't use the reference image, these algorithms can be embedded in the development of new multimedia services. NR-IQA algorithms aim at predicting image quality from objective features extracted from distorted images. To reach this goal, these algorithms, either assume a priori knowledge of involved distortions, or look for a generic approach by directly assessing the image quality regardless the type of image distortion [1]. For specific approaches, metrics are mainly designed to quantify distortions induced by image encoders such as JPEG and JPEG2000. In studies [2] [3], the block effect is estimated in the spatial domain while studies [4] [5] quantify the same effect in the frequency domain. Assessment algorithms for blur effect are proposed in [6] [7]. Distortions induced by JPEG2000 such as blur and ringing are considered in [8] [9]. The blur effect is characterized by an increase of the spread of edges while the ringing effect produces halos and/or rings near sharp object edges. As a result, the proposed metrics are generally based on the measurement of edges spreading. All of these metrics are interesting and some of them perform very well. However they remain limited by the distortions that they have to know.

Generic approaches aim to be universal and look to address all applications fields. Usually, the generic approaches follow two trends: Signal based approach that extracts and analyzes features in image signals and visually based approach that aims to mimic the human visual system (HVS) properties. Signal based approaches present a good trade-off between performance and complexity. Generally, these approaches require two steps. In the first one, relevant features are extracted while in the second, these features are pooled in order to produce the quality score of the image under test. The first step has been the subject of several investigations [10] [11] [12]. In contrast, the second step usually uses conventional combinations. To overcome this drawback, statistical modeling of natural images has been considered where the features extraction procedure is followed by a learning step. In the case of no-reference metrics which is the context of this paper, these learning methods can well map features and subjective assessments [13] [14] [15] [16] [17] [23].

Since the result of evaluation ultimately depends on the final observer, the visually-based trend looks for designing metrics according to HVS behavior when assessing quality. Several metrics based on one or more HVS properties, are proposed in literature. The HVS sensitivity to image signals is used in [24] [25]. The HVS sensitivity to spatial frequencies, luminance and/or structural information is pointed out in [26] [27] [28]. Nowadays, several HVS models are proposed in image quality assessment and their experimental results are very promising.

The goal of this paper is to combine the advantages of visually-based methods and the interesting results [29] of learning techniques in the context of NR-IQA metrics. More specifically, it uses the spatial frequency sensitivity of HVS to decompose the test image, extracts the statistical features of each visual sub-band and combines them using a multivariate Gaussian distribution, to assess the quality.

The rest of this paper is organized as follows: The proposed metric is presented in Section II. The visual sub-band decomposition is described in Section III. The experimental results are shown and discussed in section IV.

2 The LEVIQI Index

2.1 General framework

Fig. 1 gives the overall framework of the proposed method.

The first block models the spatial frequency selectivity of the human visual system and performs a perceptual decomposition on both training and testing images. The steerable pyramid [30] is used to achieve the perceptual decomposition described in Section 2.2. The steerable pyramid is a multi-scale, multi-orientation and self inverting image decomposition. With this decomposition, the image is divided into a set of sub-bands localized in scale and orientation. An example of the first level image decomposition where four oriented band-pass filters are used, is given Fig. 2.

The second block extracts the statistical features of each filtered visual channel. Instead of looking to define new features, this paper will exploit some features derived from well-known metrics.

The computed features and the DMOS (Difference of Mean Opinion Scores) values of training images are then used by the learning block to fit an IQA Multivariate Gaussian Distribution. Due to its simple parametric form, the MVGD is widely used for modeling vector-features signals and has been a sound choice in many image and video applications. Therefore, the resulting model, namely LEarning-based and Visual-based Image Quality Index (LEVIQI) is given by:

$$LEVIQI(x) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \beta)^T \Sigma^{-1} (x - \beta)\right) \quad (1)$$

where $x = (f_1, \dots, f_k, DMOS)$ corresponds to the extracted features to which is added the DMOS of training distorted images. β and Σ denote the mean and covariance matrix of the MVGD model and are estimated using the maximum likelihood method. The features extracted from testing images with DMOS values lying between 0 and 100 with a step of 0.5, are fed into the learned LEVIQI to assess objective quality of image under test. The DMOS of the test image is the one that maximizes the distribution $p(x, \beta, \Sigma)$.

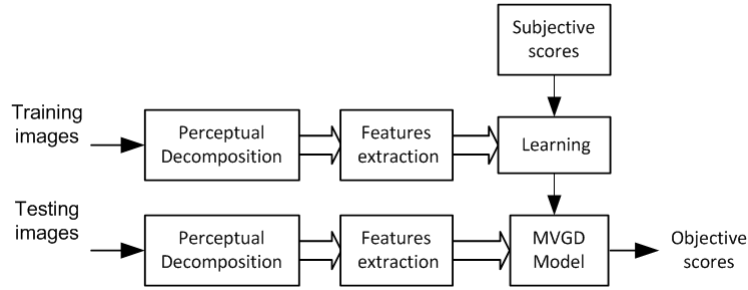


Fig. 1. Overall framework of the proposed method.

The probabilistic model is trained on the LIVE IQA database, for which, one has access to DMOS values. To ensure a robustness of results, multiple training sets were constructed. In each, the image database was subdivided into distinct training and test sets (completely content-separate). For each train set, 80% of the LIVE IQA Database content was chosen, inducing that the remaining 20% is considered for the test set. Specifically, each training set contained images derived from 23 original images, while each test set contained the images derived from the remaining 6 original images. 1000 randomly chosen training and test sets were obtained and the prediction of the quality scores was run over the 1000 iterations.

2.2 Visual Sub-band decomposition

It is well known that the retinal image is processed by different frequency channels that are narrowly tuned around specific spatial frequencies and orientations. Numerous

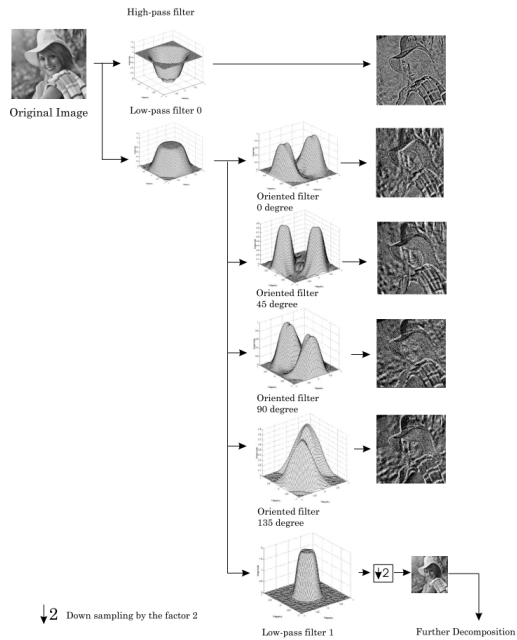


Fig. 2. Steerable pyramid-based image decomposition. Illustration of the first level image decomposition (radial selectivity) where four oriented band-pass filters are used (angular selectivity). The second level of decomposition is the result of the same process performed on low-pass filter 1 (low-pass filter 0 downsampled by the factor 2)

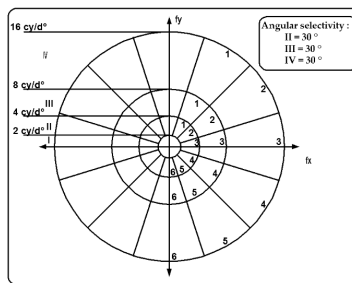


Fig. 3. Perceptual decomposition.

psychophysical experiments have been conducted to estimate the bandwidth of these channels. Quiet different values have been obtained by different types of experiments.

According to these experiments, most of the proposed decompositions suggest a spatial frequency bandwidth of approximately one octave and an orientation sensitivity that varies between 20° and 60° depending on the spatial frequency.

This study uses the decomposition of Fig. 3. Discussed in [31], this decomposition uses a radial frequency selectivity that is symmetric on a log-frequency axis with bandwidths nearly constant at one octave. It consists of one isotropic low-pass and three band-pass channels. The angular selectivity is constant and equal to 30°.

2.3 Selected features

All features considered in this paper are extracted from nine commonly used learning-based NR-IQA metrics : 1) BRISQUE [18], 2) QAC [19], 3) BLIINDS-II [17], 4) NIQE [23], 5) DIIVINE [16], 6) BIQI [15], 7) IL-NIQE [20], 8) SSEQ [21] and 9) OG-IQA [22]. A reason of the choice of those trial algorithms is motivated by the fact that the code of all of them is publicly available.

Yet, since around 200 features are available from all the trial algorithms, only the most relevant are selected. Furthermore, some of them modelize similar visual characteristics, *e.g.*, luminance sensitivity, sub-band anisotropy, and so on. So, it is not necessary to select features which modelize similar characteristics.

All features are computed for all original images (and their associated degraded version) of the LIVE IQA database [33]. Then, the Spearman Rank-Order Correlation Coefficient (SROCC) between values of features and subjective DMOS is computed. Finally, only the 10 highest correlated attributes are considered to design LEVIQI under the constraint that the selected features do not modelize similar characteristics.

| Feature from | Description | # features |
|--------------|--|------------|
| BRISQUE | shape parameter and the variance of GGD fit of the MSCN (mean subtracted contrast normalized) coefficients | 2 |
| BLIINDS2 | Coefficient of Frequency Variation | 1 |
| BLIINDS2 | Energy Subband ratio Measure | 1 |
| DIIVINE | Across scale and spatial correlations | 2 |
| SSEQ | local spatial and spectral entropy features | 2 |
| IL-NIQE | Chromatic Statistics | 2 |

Table 1. Selected features to design LEVIQI

Table 1 presents the selected distinct features. Yet, these attributes have not been used as they have been defined in their associated NR-IQA schemes, but have been modified to adapt the visual sub-band decomposition.

GGD fit parameters: The first set of features, derived from BRISQUE, includes the shape parameter α and the variance σ^2 of generalized Gaussian distribution (GGD) fit of the MSCN (mean subtracted contrast normalized) coefficients for each sub-band. The MSCN coefficients refer to the transformed luminances $\hat{I}(i, j)$ given by:

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C}, \quad (2)$$

where (i,j) are spatial indices, $C=1$, and $\mu(i,j) = \sum_{k=-3}^3 \sum_{l=-3}^3 \omega_{k,l} I_{k,l}(i,j)$, and $\sigma(i,j) = \sqrt{\sum_{k=-3}^3 \sum_{l=-3}^3 \omega_{k,l} (I_{k,l}(i,j) - \mu(i,j))^2}$, ω is a 2D circularly-symmetric Gaussian weighting function.

The shape α and the variance σ^2 parameter of the GGD are computed over all subbands and then pooled by computing the lowest 10th percentile average of the local α scores and the local σ^2 scores across the sub-bands.

Coefficient of Frequency Variation: This feature, derived from BLIINDS2, is defined for each subband as:

$$\zeta = \frac{\sigma_{|X|}}{\mu_{|X|}} = \sqrt{\frac{\Gamma(1/\gamma) \Gamma(3/\gamma)}{\Gamma^2(2/\gamma)} - 1} \quad (3)$$

where $\sigma_{|X|}$ and $\mu_{|X|}$ are the standard deviation and mean of the frequency coefficient magnitudes $|X|$, respectively.

The feature ζ is computed for all sub-bands of the image and pooled by taking the highest 10th percentile and over all of the sub-band scores of the image.

Energy Sub-band ratio Measure: This attribute, also derived from BLIINDS2, is used to capture local spectral signature. This feature is defined as:

$$R_{n,k} = \frac{\left| E_{n,k} - \frac{1}{n-1} \sum_{j < n} E_{j,k} \right|}{E_{n,k} + \frac{1}{n-1} \sum_{j < n} E_{j,k}} \quad (4)$$

where $E_{n,k} = \sigma_{n,k}^2$ is the average energy in the frequency sub-band n for a given radial band k , where $n \in \{1, 2, 3, \dots, 6\}$ and $k \in \{1, 2, 3\}$. The mean of $(R_{n,k})_{n \in [2, \dots, 6]}$ is computed on the three radial bands k of the image and pooled by computing the highest 10th percentile average of the scores of the image.

Across scale and Spatial correlations These two features are derived from features designed for DIIVINE. In order to capture the statistical dependencies between high-pass (HP) responses of natural images and their band-pass (BP) counterparts, a structural correlation is modeled as $\rho = (2\sigma_{xy} + C_2)/(\sigma_x^2 + \sigma_y^2 + C_2)$ where σ_{xy} is the cross-variance between the windowed regions from the BP and HP bands, and σ_x^2 , σ_y^2 are their windowed variances respectively; C_2 is a stabilizing constant. The mean of the 18 correlation values (corresponding to the 18 sub-bands) is computed.

In order to capture spatial correlation statistics, the joint empirical distribution between coefficients at (i,j) and the set of spatial locations at a distance of τ is computed for each d_1^θ , $\theta \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$. The correlation between these two variables denoted X and Y is estimated by:

$$\rho(\tau) = \frac{E_{P_{XY}(x,y)} \left[(X - E_{P_X(x)}[X])^T (Y - E_{P_Y(y)}[Y]) \right]}{\sigma_X \sigma_Y} \quad (5)$$

where $E_{P_X(x)}[X]$ is the expectation of X with respect to the marginal distribution $p_X(x)$ (similarly for Y and (X,Y)). $\rho(\tau)$ is plotted for various distance across distortions and the obtained curve is fitted with a 3rd order polynomial. The coefficients of the polynomial and the error between the fit and the actual $\rho(\tau)$ form an 30 dimensional feature vector (5 values for each direction). The mean of $\rho(\tau)$ is computed over the three radial bands and pooled by computing the highest 10th percentile average of the scores.

Local spatial and spectral entropy features: The spatial entropy is computed on the obtained image after applying the first non-directional lower-pass filter applying.

Concerning the spectral entropy, a modified version of the attribute designed for SSEQ is defined. The variance across the six orientations per radial band $var(E[R_\theta])$ is computed. ($E[R_\theta]$) is the average of the Renyi entropy $R_\theta[n]$ per orientation for the three sub-bands (one per radial band) of orientation θ .

Finally, the mean of the three variance values per radial band is computed.

Chromatic Statistics: This attributed is derived from IL-NIQE and is computed on the color image before applying the DCP transform. From the RGB coordinates system, a logarithmic-scale opponent color space \mathcal{RGB} is defined as $\mathcal{R} = \log R - \mu_R$, $\mathcal{G} = \log G - \mu_G$ and $\mathcal{B} = \log B - \mu_B$

where μ_R (resp. μ_G and μ_B) is the mean $\log R$ (resp. $\log G$ and $\log B$) over the image. Finally, the following linear color transform is applied on the \mathcal{RGB} color space as $l_1 = (\mathcal{R} + \mathcal{G} + \mathcal{B})/\sqrt{3}$, $l_2 = (\mathcal{R} + \mathcal{G} - 2\mathcal{B})/\sqrt{6}$, $l_3 = (\mathcal{R} - \mathcal{G})/\sqrt{2}$.

The distributions of each channel l_1, l_2, l_3 conforms to a Gaussian probability law. In this paper, only the chromatic channels l_2 and l_3 are considered, since l_1 refers to a luminance channel. Finally, the two model parameters μ_C and σ_C^2 are estimated using a multivariate Gaussian model.

3 Performance evaluation

3.1 Experimental setup

Trial databases: To provide comparison of NR-IQA algorithms, two publicly available databases are used: 1) TID2013 database [32] and 2) CSIQ database [34]. Since the LIVE database [33] has been used to train both the proposed metric and most of the trail NR-IQA schemes, it has not been used to evaluate performances of NR-IQA methods.

The TID2013 database contains images with multiple distortions. This database consists of 25 original images on which 24 different type of distortions have been applied using five degradation levels per distortion. A total of 3000 distorted images are generated. It is worth noting that seven new types of degradations have been introduced with respect to the 17 types of degradation existing in the previous version of the database, known as TID2008. The database contains 524340 subjective ratings from 971 different observers, and ratings are reported in the form of MOS.

The CSIQ Database consists of 30 original images, each is distorted using six different types of distortions at four to five different levels of distortion. A total of 866 distorted images have been generated. The database contains 5000 subjective ratings from 35 different observers, and ratings are reported in the form of DMOS.

| TID2013 subset | BRISQUE | BLIINDS2 | DIIVINE | SSEQ | ILNIQE | LEVIQI |
|------------------------------------|--------------|----------|--------------|--------------|--------------|--------------|
| Additive Gaussian Noise | 0.852 | 0.722 | 0.855 | 0.807 | 0.876 | 0.881 |
| Additive Noise in Color Components | 0.709 | 0.649 | 0.712 | 0.681 | 0.815 | 0.817 |
| Spatially Correlated Noise | 0.491 | 0.767 | 0.463 | 0.635 | 0.923 | 0.899 |
| Masked Noise | 0.575 | 0.512 | 0.675 | 0.565 | 0.512 | 0.779 |
| High Frequency Noise | 0.753 | 0.824 | 0.878 | 0.860 | 0.868 | 0.891 |
| Impulse Noise | 0.630 | 0.650 | 0.806 | 0.749 | 0.755 | 0.801 |
| Quantization Noise | 0.798 | 0.781 | 0.165 | 0.468 | 0.873 | 0.860 |
| Gaussian Blur | 0.813 | 0.855 | 0.834 | 0.858 | 0.814 | 0.851 |
| Image Denoising | 0.586 | 0.711 | 0.723 | 0.783 | 0.750 | 0.812 |
| JPEG Compression | 0.852 | 0.864 | 0.629 | 0.825 | 0.834 | 0.867 |
| JPEG2000 Compression | 0.893 | 0.898 | 0.853 | 0.885 | 0.857 | 0.901 |
| JPEG Transmission Errors | 0.315 | 0.117 | 0.239 | 0.354 | 0.282 | 0.217 |
| JPEG2000 Transmission Errors | 0.360 | 0.620 | 0.060 | 0.561 | 0.524 | 0.662 |
| Non Eccentricity Pattern Noise | 0.145 | 0.096 | 0.060 | 0.011 | 0.080 | 0.121 |
| Local Block-wise Distortions | 0.224 | 0.209 | 0.093 | 0.016 | 0.135 | 0.217 |
| Mean Shift | 0.124 | 0.128 | 0.010 | 0.108 | 0.184 | 0.211 |
| Contrast Change | 0.040 | 0.150 | 0.460 | 0.204 | 0.014 | 0.521 |
| Change of Color Saturation | 0.109 | 0.017 | 0.068 | 0.074 | 0.162 | 0.471 |
| Multiplicative Gaussian Noise | 0.724 | 0.716 | 0.787 | 0.679 | 0.693 | 0.771 |
| Comfort Noise | 0.008 | 0.017 | 0.116 | 0.033 | 0.359 | 0.315 |
| Lossy Compression of Noisy Images | 0.685 | 0.719 | 0.633 | 0.610 | 0.828 | 0.831 |
| Color Quantization with Dither | 0.764 | 0.736 | 0.436 | 0.528 | 0.748 | 0.892 |
| Chromatic Aberrations | 0.616 | 0.539 | 0.661 | 0.688 | 0.679 | 0.699 |
| Sparse Sampling and Reconstruction | 0.784 | 0.816 | 0.834 | 0.895 | 0.865 | 0.899 |
| Cumulative subsets | 0.367 | 0.393 | 0.355 | 0.332 | 0.494 | 0.501 |

Table 2. SROCC values computed between predicted scores using NR-IQA schemes from which some attributes are extracted to design LEVIQI and MOS values for TID2013 Images database.

Trial NR-IQA: To assess the performance of the proposed metric, nine commonly used NR-IQA schemes are used to compare LEVIQI with. The trial metrics are those mentioned in section 2.3. Only NR-IQA schemes whose at least one feature has been adapted to design LEVIQI index, *i.e.*, BRISQUE, BLIINDS2, DIIVINE, SSEQ and ILNIQE are used to compare the performance of LEVIQI with.

Statistical Significance and hypothesis testing: Results obtained from the proposed metric are compared to results provided by all trial NR-IQA algorithms. To perform this evaluation, the Spearman Rank Order Correlation Coefficient (SROCC) is computed between the DMOS values and the predicted scores obtained from NR-IQA algorithms. In addition, to ascertain which differences between NR-IQA schemes performance are statistically significant, we applied an hypothesis test using the residuals between the DMOS values and the ratings provided by the IQA algorithms. This test is based on the t-test that determines whether two population means are equal or not. This test yields us to take a statistically-based conclusion of superiority (or not) of an NR-IQA algorithm.

3.2 Experimental results

Table 2 gives SROCC mean values computed between predicted scores from NR-IQA schemes from which some extracted features are used to design LEVIQI and MOS values for the TID2013 Images database. When considering the whole database, LEVIQI

overperforms BRISQUE, BLIINDS2, DIIVINE, SSEQ ad ILNIQE. In 67% of cases (16 subsets out of 24), scores predicted by LEVIQI allow to have a better correlation (higher SROCC value) with human judgments than that of any other tested NR-IQA. For the remaining cases (8 remaining subsets), the performance of LEVIQI is very close to the best values. If we consider multidistorted images associated to the seven last subset of Table 2, LEVIQI performs better than trial quality schemes except for 'Multiplicative Gaussian Noise' and 'Comfort Noise'. Yet, obtained SROCC values are highly competitive with the best quality index.

Similar results as those of Table 2 are given in Table 3 for CSIQ Images database. For this database also, The global performance (SROCC value of cumulative subsets) of LEVIQI is higher than that of trial metrics. More specifically, LEVIQI performs better 4 times over 6. For the two remaining distortions 'Gaussian Noise' and 'Additive Gaussian Pink Noise', LEVIQI presents the second best performance.

| CSIQ subset | BRISQUE | BLIINDS2 | DIIVINE | SSEQ | ILNIQE | LEVIQI |
|---------------------------|--------------|----------|---------|--------------|--------|--------------|
| JP2K | 0.866 | 0.895 | 0.830 | 0.848 | 0.906 | 0.911 |
| JPEG | 0.903 | 0.901 | 0.799 | 0.865 | 0.899 | 0.915 |
| Gaussian Noise | 0.252 | 0.379 | 0.176 | 0.872 | 0.850 | 0.862 |
| Add. Gaussian Pink Noise | 0.925 | 0.801 | 0.866 | 0.046 | 0.874 | 0.909 |
| Gaussian Blur | 0.903 | 0.891 | 0.871 | 0.873 | 0.858 | 0.931 |
| Global Contrast Decrement | 0.029 | 0.012 | 0.396 | 0.200 | 0.501 | 0.590 |
| Cumulative subsets | 0.566 | 0.577 | 0.596 | 0.528 | 0.815 | 0.821 |

Table 3. SROCC values computed between predicted scores from which some attributes are extracted to design LEVIQI and MOS values for CSIQ Images database.

In addition, Table 4 gives obtained results when a One-sided t-test is used to provide statistical significance of NR-IQA/LEVIQI quality scores on the 6 multidistortions subsets of TID2013 database (the 6 last subsets in Table 2). Each entry in this table is coded using six symbols. The position of each symbol corresponds to one subset (first position corresponds to 'Change of Color Saturation' subset, second position for 'Multiplicative Gaussian Noise' subset etc..). Each symbol gives the result of the hypothesis test on the subset. If the symbol equals '1', the NR-IQA on the row is statistically better than the NR-IQA on the column ('0' means worse, '-' is used when NR-IQAs are indistinguishable). Results confirm that, most of the time, LEVIQI is more consistent with human judgments than trial NR-IQA schemes from which some features have been extracted. As multidistortions subsets are not common to LIVE database, these results are more reliable.

| | BRISQUE | BLIINDS2 | DIIVINE | SSEQ | ILNIQE |
|--------|---------|----------|---------|---------|---------|
| LEVIQI | 1111111 | 1111111 | 10111-1 | 11111-- | 110-1-- |

Table 4. Statistical significance comparison of trial NR-IQA/LEVIQI quality scores on TID2013 database multidistortions subsets.

In a similar way, Table 5 gives obtained results when a One-sided t-test is used to provide statistical significance of NR-IQA/LEVIQI quality scores on CSIQ database. One can notice that distortions present in CSIQ database are also present in LIVE database. For these learned distortions, LEVIQI exceeds in several cases all of the trial NR-IQA schemes.

| | BRISQUE | BLIINDS2 | DIIVINE | SSEQ | ILNIQE |
|--------|---------|----------|---------|---------|---------|
| LEVIQI | 1-1-111 | --11111 | 1111111 | 1101111 | ---111- |

Table 5. Statistical significance comparison of NR-IQA/LEVIQI quality scores on CSIQ database subsets.

Finally, to compare the computational complexity of the proposed algorithm, we measured the average computation time required to assess an image of size 512×512 , using a computer with Intel Core-I7 processor at 2.2GHz. Table 6 reports the measurement results, which are rough estimates only, as no code optimization has been done on our Matlab implementations. LEVIQI is superior to DIIVINE and BLIINDS2 while inferior to BRISQUE and SSEQ and similar to ILNIQE.

| NR-IQA | BRISQUE | BLIINDS2 | DIIVINE | SSEQ | ILNIQE | LEVIQI |
|--------|---------|----------|---------|------|--------|--------|
| time | 1.33 | 87.13 | 27.33 | 2.69 | 11.85 | 12.38 |

Table 6. Comparison of computational time (in second/image)

4 Conclusion

In this paper, a machine learning-based and human vision-based quality index called LEVIQI has been proposed in the purpose of no reference quality evaluation. The model utilizes ten derivative relevant attributes from five well-known and highly competitive NR-IQA schemes. The selected attributes address human vision characteristics such as frequency sensitivity, chromatic sensitivity, anisotropy and contrast sensitivity. These attributes have been adapted to be computed on 18 sub-bands generated from three different radial bands and six different orientations, to simulate human perception sensitivity. Obtained results show that the performance of LEVIQI is competitive with top-performing NR-IQA schemes. The potential of this model to be used in real applications is the subject of investigations where real-time implementation of LEVIQI index is considered using efficient vectorization.

References

1. Zhang, S., Zhang, C., Zhang, T., Lei, Z.: Review on universal no-reference image quality assessment algorithm. *Comput. Eng. Appl.* 51, 13–23 (2015)
2. Gao, W., Mermer, C., and Kim, Y.: A de-blocking algorithm and a blockiness metric for highly compressed images. *IEEE trans. Circuits and Syst. for Video Tech.*, 12(12), 1150–1159 (2002)
3. Pan, F., Lin, X., Rahardja, S., Lin, W., Ong, E., Yao, S., Lu, Z., and Yang, X.: A locally adaptive algorithm for measuring blocking artifacts in images and videos. *Signal Processing: Image Communication*, 19(6), 499–506 (2004)
4. Wang, Z., Bovik, A. C., and Evan, B.: Blind measurement of blocking artifacts in images. In: *International Conference on Image Processing*, volume 3, pp. 981–984. IEEE Press, (2000).
5. Liu, S. and Bovik, A. C.: Efficient dct-domain blind measurement and reduction of blocking artifacts. In: *IEEE Trans. Circuits and Syst. for Video Tech.*, 12(12), 1139–1149 (2002)
6. Ferzli, R. and Karam, L.J.: A no-reference objective image sharpness metric based on the notion of Just Noticeable Blur (JNB). *IEEE Trans. Image Process.* 18, 717–728 (2009)
7. Sang, Q.B.; Su, Y.Y.; Li, C.F.; Wu, X.J.: No-reference blur image quality assessment based on gradient similarity. *J. Optoelectron. Laser* 24, 573-577 (2013)
8. Sazzad, Z.M.P., Kawayoke, Y., Horita, Y.: No reference image quality assessment for JPEG2000 based on spatial features. *Signal Process. Image Commun.* 23, 257-268 (2008)
9. Barland, R., Saadane, A.: Reference free quality metric for JPEG-2000 compressed images. In: *8th Inter. Symp. on Signal Processing and Its Applications*, pp. 351–354. (2005)

10. Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P.: Image quality assessment: From error measurement to structural similarity. *IEEE Trans. Im. Proc.* 13(1), 996-1006 (2004)
11. Shnayderman, A., Gusev, A., Eskicioglu, A.: An svd-based grayscale image quality measure for local and global assessment. *IEEE Trans. On Image Processing.* 15, 2, pp. 422-429 (2006)
12. Wee, C., Paramesran, R., Mukundan, R., Jiang, X.: Image quality assessment by discrete orthogonal moments. *Pattern Recognition.* 43, 12, pp. 4055-4068 (2010)
13. Babu, R. V., Suresh, S., Perki, A.: No-reference JPEG image quality assessment using GAP-RBF. *Signal Processing.* 87 (6), 1493-1503 (2007)
14. Gabarda, S., Cristobal, G.: Blind image quality assessment through anisotropy. *JOSA.* 24 (12), B42-B51 (2007)
15. Moorthy, A. K., Bovik, A. C.: A two-step framework for constructing blind image quality indice. *IEEE Signal processing letters.* 17 (5), 513-516 (2010)
16. Moorthy, A. K., Bovik, A. C.: Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing.* 20 (12), 3350-3364 (2011)
17. Saad, M., Bovik, A. C., Charrier, C.: Blind image quality assessment: A natural scene statistics approach in the dct domain. *IEEE Trans. on Image Processing.* 21 (8), 3339-3352 (2012)
18. A. Mittal, A. K. Moorthy, A. C. Bovik: No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, 2012.
19. Xue, W., Zhang, L., Mou, X.: Learning without human scores for blind image quality assessment. *IEEE Conf. Proc. Vis. pattern Recog.*, pp 995-1002, 2013.
20. L. Zhang, L. Zhang, A. C. Bovik: A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579-2591, 2015.
21. L. Liu, B. Liu, H. Huang, A.C. Bovik: No-reference image quality assessment based on spatial and spectral entropies. *Sig. Proc.: Image Com.*, vol. 29, pp. 856-863, 2014.
22. L. Liu, Y. Hua, Q. Zhao, H. Huang, A. C. Bovik: Blind image quality assessment by relative gradient statistics and ada boosting neural network. *Sig. Proc.: Img. Com.*, vol. 40, 1-15, 2016.
23. Mittal, A., Soundararajan, R., Bovik, A. C.: Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* 20 (3), 209-212 (2013)
24. Chandler, D.M., Hemami, S.S.: VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing.* 16(9), 2284-2298 (2007)
25. Sheikh, H.R. Bovik, A.C.: Image information and visual quality. *IEEE Transactions on Image Processing.* 15(2), 430-444 (2006).
26. Larson, E.C., Chandler, D.M.: Most apparent distortion: Full-reference image quality assessment and the role of strategy. *J. Electron. Imag.* 19(1), 011006-1-011006-21 (2010)
27. Pei, S.C., Chen, L.H.: Image Quality Assessment Using Human Visual DOG Model Fused With Random Forest. *IEEE Transactions on Image Processing.* 24(11), 3282-3292 (2015)
28. Le Callet, P., Saadane, A., Barba, D.: Frequency and spatial pooling of visual differences for still image quality assessment. In *Proc. SPIE*, vol. 3959, pp. 595-603 (2000).
29. Gastaldo, P., Zunino, R., Redi, J.: Supporting visual quality assessment with machine learning. *EURASIP J. Image Video Process.* , 2013:54 (2013).
30. Liu, Z., Tsukada, K., Hanasaki, K., Ho, Y.K., Dai, Y.P.: Image fusion by using steerable pyramid. *Pattern Recognition Letters?* 22, 929-939 (2001)
31. Daly, S.: A visual model of optimizing the design of image processing algorithm. In *Proc. ICIP*, Vol. II of III, pp. 16-20 (1994).
32. N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo: Image database TID2013: Peculiarities, results and perspectives. *Signal Processing: Image Communication*, vol. 30, Jan. 2015, pp. 57-77.
33. Laboratory for Image & Video Engineering, University of Texas (Austin), LIVE Image Quality Assessment Database, <http://live.ece.utexas.edu/research/Quality>, 2002.
34. E. C. Larson, D. M. Chandler: Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging* 19(1), pp. 1-21, 2010.