

RESEARCH

Open Access



Referenceless image quality assessment by saliency, color-texture energy, and gradient boosting machines

Pedro Garcia Freitas^{1*} , Wellington Y. L. Akamine² and Mylène C. Q. Farias²

Abstract

In most practical multimedia applications, processes are used to manipulate the image content. These processes include compression, transmission, or restoration techniques, which often create distortions that may be visible to human subjects. The design of algorithms that can estimate the visual similarity between a distorted image and its non-distorted version, as perceived by a human viewer, can lead to significant improvements in these processes. Therefore, over the last decades, researchers have been developing quality metrics (i.e., algorithms) that estimate the quality of images in multimedia applications. These metrics can make use of either the full pristine content (full-reference metrics) or only of the distorted image (referenceless metric). This paper introduces a novel referenceless image quality assessment (RIQA) metric, which provides significant improvements when compared to other state-of-the-art methods. The proposed method combines statistics of the opposite color local variance pattern (OC-LVP) descriptor with statistics of the opposite color local salient pattern (OC-LSP) descriptor. Both OC-LVP and OC-LSP descriptors, which are proposed in this paper, are extensions of the opposite color local binary pattern (OC-LBP) operator. Statistics of these operators generate features that are mapped into subjective quality scores using a machine-learning approach. Specifically, to fit a predictive model, features are used as input to a gradient boosting machine (GBM). Results show that the proposed method is robust and accurate, outperforming other state-of-the-art RIQA methods.

Keywords: Color-texture analysis, Gradient boosting machine, Local variance patterns, Local salient patterns, Referenceless image quality assessment, Regression, Visual attention

Background

The rapid growth of current multimedia industry, and the consequent increase in content quality requirements, have prompted the interest in visual quality assessment methodologies [1]. Because most multimedia applications are designed for human observers, visual perception has to be considered when measuring visual quality [2]. Psychophysical experiments (or subjective quality assessment methods) performed with human subjects are considered the most accurate methods to assess visual quality [3]. However, these subjective methods are costly, time-consuming, and, for this reason, not adequate for real-time multimedia applications.

Objective quality assessment metrics predict visual quality employing mathematical methods instead of human subjects. For instance, mean squared deviation (MSD) and peak-to-noise ratio (PSNR) are mathematical methods that can be used to measure the similarity of visual signals. However, MSD and PSNR scores often do not correlate well with the image quality as perceived by human observers (i.e., subjective scores) [4]. It is worth mentioning that, for an objective metric to be used in multimedia applications, its estimates must be well correlated with quality scores available in publicly available quality databases, which use standardized experimental procedures to measure the quality of a comprehensive series of visual signals.

Metrics can be classified according to the quantity of reference information (pristine content) required by the method. While full-reference (FR) metrics require the original content, reduced-reference (RR) metrics demand

*Correspondence: pedrogarcia@ieee.org

¹Department of Computer Science, University of Brasília, Campus Universitário Darcy Ribeiro, Brasília - DF 70919-970 Brazil

Full list of author information is available at the end of the article

only parts of original information. Since the reference (or even partial reference information) is not available in many multimedia applications, there is a need for referenceless metrics that do not require any information about the reference image.

The development of referenceless image quality assessment (RIQA) methods remains a challenging problem [2, 5]. A popular approach consists of estimating image quality using distortion-specific (DS) methods that measure the intensity of the most relevant image distortions. Among the state-of-the-art DS methods, we can cite the papers of Fang et al. [6], Bahrami and Kot [7], Golestaneh and Chandler [8], and Li et al. [9–11]. These methods make assumptions about the type of distortion present in the signal and, as a consequence, have limited applications in more diverse multimedia scenarios.

Non-distortion-specific (NDS) methods, which do not demand a prior knowledge about the type of distortions in the signal, are more suitable for diverse multimedia scenarios. In this case, instead of making assumptions about the main characteristics of specific distortions, the methods make assumptions about the image characteristics. For instance, to find the relationship between gradient information and image quality, Liu et al. [12] and Li et al. [13] make assumptions about the image structure of reference images in the gradient domain. Some methods compare the statistics of impaired and non-impaired (natural) images using a “natural scene statistic” (NSS) approach [14, 15].

In addition to the aforementioned approaches, IQA methods can be classified as feature-based or human visual system (HVS)-based approaches. Feature-based approaches extract and analyze features from image signals to estimate quality. Usually, these approaches require three steps. In the first step, descriptive features are extracted. Then, the extracted features are pooled to produce a quality-aware feature vector. Finally, a model maps the pooled data into a numerical value that represents the quality score of the image under test. One example of a feature-based metric is the work of Mittal et al. [16], which is a spatial-domain method based on the NSS. Saad et al. [14, 17] proposed another feature-based NSS method that operates in the discrete cosine transform (DCT) domain. Finally, Liu et al. [18] proposed a feature-based method that is based on spatial and spectral image entropies. More recently, some works proposed feature extraction used texture information to estimate image quality [19–27].

Instead of extracting basic features from images, HVS-based approaches aim to mimic the HVS behavior. Hitherto, various HVS properties have been used in quality metrics, including structural information [28, 29] and error and brightness sensitivities [30, 31]. The acclaimed structural similarity index (SSIM) [32] is based on the assumption that HVS is more sensitive to the structural

information of the visual content and, therefore, a structural similarity measure can provide a good estimate of the perceived image quality. The recent free energy theory revealed that the HVS strives to comprehend the input visual signal by reducing the undetermined portions, which affects the perception of quality [33]. Zhang et al. [34] proposed a Riesz transform-based feature similarity index (RFSIM) that characterizes local structures of images and uses a Canny edge detector to generate a pooling mask. More recently, HVS-based methods employing convolutional neural networks (CNN) have been proposed [35–37]. These CNN-based methods are established on the comparison between the hierarchy of the human visual areas and the layers of a CNN [38, 39].

In recent years, HVS-based image quality approaches that incorporate visual saliency models (VSM) have been a trend [40–43]. Image quality metrics and VSM are inherently correlated because both of them take into account how the HVS perceives the visual content (i.e., how humans perceive suprathreshold distortions) [42]. Since VSM provide a measurement of the region’s importance, they can be successfully used for weight distortions in image quality algorithms. Several researchers have studied how the saliency information can be incorporated into visual quality metrics to enhance their performance [41, 44–47]. However, most VSM-based quality metrics are FR approaches. Among the existing VSM-based RIQA methods, most are DS methods that cannot be used as general-purpose RIQA methods (GP-RIQA).

Additionally, most of current GP-IQA methods have no good prediction accuracy for color and contrast-distorted images. For instance, Ortiz-Jaramill et al. [48] demonstrated that current color difference measures (i.e., FR-IQA methods that compute color differences between processed and reference images) present little correlation with subjective quality scores. Also, even though some DS-IQA methods are able to predict the quality of contrast-distorted images [49], most GP-IQA methods have a poor prediction performance. This low performance leads to authors often omitting the results for these types of image distortions [18, 20, 23, 50].

In this paper, we introduce a NDS-GP-RIQA method based on machine learning (ML) that tackles these limitations by taking into account how impairments affect salient color-texture and energy information. The introduced method is based on the statistics of two new proposed descriptors: the opponent color local variance pattern (OC-LVP) and the opposite-color local salient pattern (OC-LSP). These proposed descriptors are extensions of the opponent-color local binary pattern (OC-LBP) [51] that incorporate both feature-based and HVS-based approaches. More specifically, the OC-LSP extends the OC-LBP by encoding both spatial, color, and

saliency information using a VSM to weight the OC-LBP statistics. The OC-LVP descriptor uses concepts introduced by the local variant patterns (LVP) [52] to modify the OC-LBP and measure the color-texture energy. The method uses the statistics of OC-LVP and OC-LSP as input of a gradient boosting machine (GBM) [53, 54] that learns the predictive quality model via regression. When compared to previous work [52], in this work, we use OC-LSP and OC-LVP operators, instead of the simpler LVP operator. The metric design of the metric was also modified to use a GBM, instead of the random forest regression algorithm.

The rest of this paper is divided as follows. In the “A brief review of local binary patterns” section, the basis of texture analysis is revised. In the “Opponent color local binary pattern” section, the base color-texture descriptor is summarized. In the “Opponent color local salient pattern” and “Opposite color local variance pattern” sections, the proposed descriptors are detailed. In the “Feature extraction” and “Gradient boosting machine for regression” sections, we describe how to use the proposed descriptors to predict image quality without references. An extensive analysis of the results is presented in the “Results and discussion” section. Finally, the “Conclusions” section concludes this paper.

Methods

In this section, we review the basic texture operator local binary pattern (LBP) and its improved color-texture extension, the opposite color local binary patterns (OC-LBP). Then, we describe the proposed quality-aware descriptor, named the color local salient patterns (OC-LSP) and the color local variance patterns (OC-LVP). Finally, this section finishes with the proposed quality assessment method based on these operators.

A brief review of local binary patterns

Local binary pattern (LBP) is indubitably one of the most effective texture descriptors available for texture analysis of digital images. It was first proposed by Ojala et al. [55] as a specific case of the texture spectrum model [56]. Being $I \in \mathbb{R}^{m \times n}$ the image whose texture we want to describe, the ordinary LBP takes the form:

$$LBP_{R,P}(I_c) = \sum_{p=0}^{P-1} f(I_p, I_c, p), \tag{1}$$

where

$$f(I_p, I_c, p) = S(I_p - I_c) \cdot 2^p \tag{2}$$

and

$$S(t) = \begin{cases} 1, & \text{if } t \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

In Eq. 1, $I_c = I(x, y)$ is an arbitrary central pixel at the position (x, y) and $I_p = I(x_p, y_p)$ is a neighboring pixel surrounding I_c , where:

$$x_p = x + R \cos\left(2\pi \cdot \frac{p}{P}\right),$$

and

$$y_p = y - R \sin\left(2\pi \cdot \frac{p}{P}\right).$$

In this case, P is the number of neighboring pixels sampled from a distance of R from I_c to I_p . Figure 1 illustrates examples of symmetric samplings for different neighboring points (P) and radius (R) values.

Figure 2 exemplifies the steps for applying the LBP operator on a single pixel ($I_c = 35$), located in the center of a 3×3 image block, as shown in the bottom-left of this figure. The numbers in the yellow squares of the block represent the order in which the operator is computed (counter-clockwise direction starting from 0). In this figure, we use an unitary neighborhood radius ($R = 1$) and eight neighboring pixels ($P = 8$). After calculating $S(t)$ (see Eq. 3) for each neighboring pixel I_p , we obtain a binary output for each I_p ($0 \leq p \leq 7$), as illustrated in the block in the upper-left position of Fig. 2. In this block, black circles correspond to “0” and white circles to “1”. These binary outputs are stored in a binary format, according to their position (yellow squares). Then, the resulting binary number is converted to the decimal format. For a complete image, we use the LBP operator to obtain a decimal number for each pixel of the image, by making I_c equal the current pixel.

When an image is rotated, I_p values move along the perimeter of the circumference (around I_c), generating a

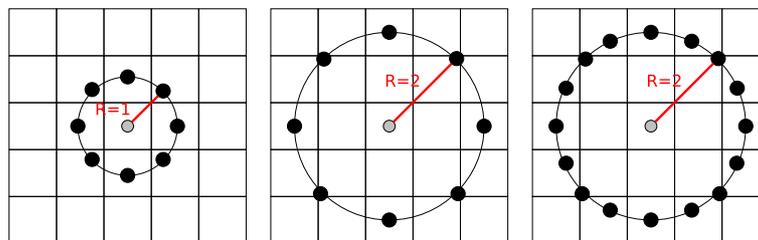
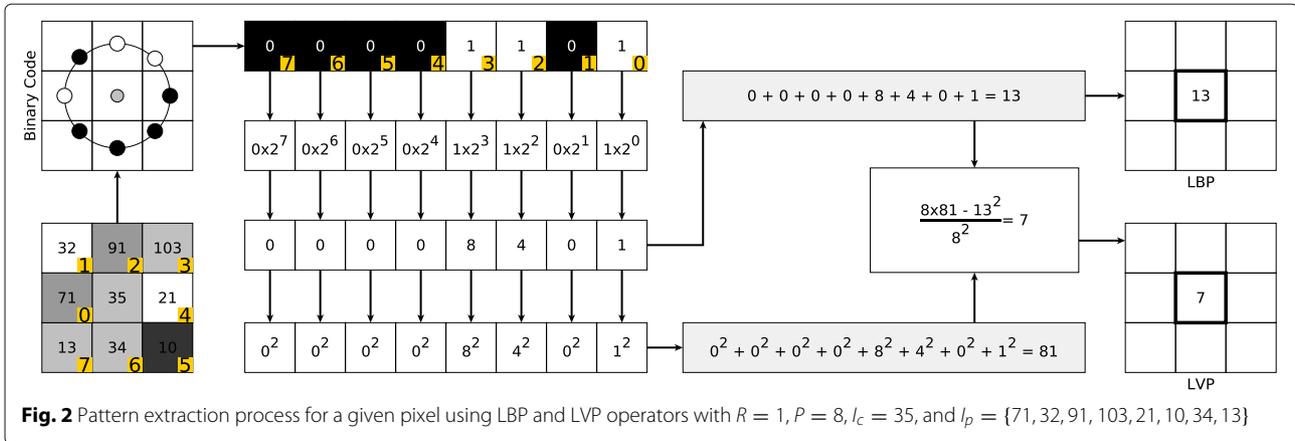


Fig. 1 Circularly symmetric P neighbors extracted from a distance R



circular shift in the binary number generated. As a consequence, a different decimal $LBP_{R,P}(I_c)$ value is obtained. To remove this effect, we assign a unique identifier to each rotation, generating a rotation invariant LBP:

$$LBP_{R,P}^i(I_c) = \min \{ ROTR(LBP_{R,P}(I_c), k) \}, \quad (4)$$

where $k = \{0, 1, 2, \dots, P - 1\}$ and $ROTR(x, k)$ is the circular bit-wise right shift operator that shifts the t-tuple x by moving k positions.

Due to the primitive quantization of the angular space [57, 58], $LBP_{R,P}$ and $LBP_{R,P}^i$ operators do not always provide a good discrimination [58]. To improve the discriminability of the LBP operator, Ojala et al. [55] proposed an improved operator that captures fundamental pattern properties. These fundamental patterns are called “uniform” and computed as follows:

$$LBP_{R,P}^u(I_c) = \begin{cases} \sum_{p=0}^{P-1} f(I_p, I_c, p) & U(LBP_{R,P}^i) \leq 2, \\ P + 1 & \text{otherwise,} \end{cases} \quad (5)$$

where $U(LBP_{P,R})$ is the uniform pattern given by:

$$U(LBP_{P,R}) = \Delta(I_{P-1}, I_0) + \sum_{p=1}^{P-1} \Delta(I_p, I_{p-1}), \quad (6)$$

and

$$\Delta(I_x, I_y) = |S(I_x - I_c) - S(I_y - I_c)|. \quad (7)$$

In addition to a better discriminability, the uniform LBP operator (Eq. 5) has the advantage of generating fewer distinct LBP labels. While the “nonuniform” operator (Eq. 1) produces 2^P different output values, the uniform operator produces only $P + 2$ distinct output values, and the “rotation invariant” operator produces $P(P - 1) + 2$ points.

Opponent color local binary pattern

The LBP operator is designed to characterize texture of grayscale images. Although this restriction may not affect

many applications, it may be unfavorable for image quality assessment purposes because LBP is not sensitive to some types of impairments, such as contrast distortions or chromatic aberrations. As pointed out by Maenpaa et al. [51], texture and color have interdependent roles. When luminance-based texture descriptors (e.g., LBP) achieve good results, color descriptors can also obtain good results. However, when color descriptors are unsuccessful, luminance texture descriptors can still present a good performance. For this reason, operators that integrate both color and texture information tend to be more successful to predict the quality of images with a wider range of distortions.

In order to integrate color and texture into a single descriptor, Maenpaa et al. [51] introduced the opponent color local binary pattern (OC-LBP). The OC-LBP extends the LBP operator by incorporating color information, while keeping texture information. This color-texture descriptor is an extension of the operator proposed by Jain and Healey [59], which replaces the Gabor’s filtering with a variant of the LBP-inspired operator.

The OC-LBP descriptor operates on intra-channel and inter-channel color dimensions. In the intra-channel operation, the LBP operator is applied individually, on each color channel, instead of being applied only on a single luminance channel. This approach is called “intra-channel” because the central pixel and the corresponding sampled neighboring points belong to the same color channels.

In the “inter-channel” operation, the central pixel belongs to a color channel and its corresponding neighboring points are necessarily sampled from another color channel. Therefore, for a three-channel color space, such as HSV, there are six possible combinations of channels: OC-LBP_{HS}, OC-LBP_{SH}, OC-LBP_{HV}, OC-LBP_{VH}, OC-LBP_{SV}, and OC-LBP_{VS}.

Figure 3 illustrates the sampling approach of OC-LBP when the central pixel is sampled in the R channel of a

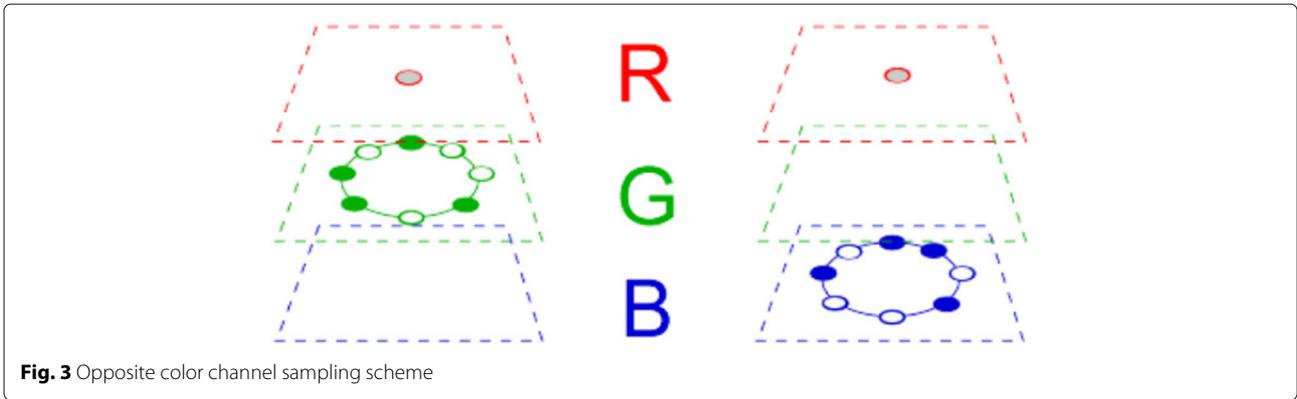


Fig. 3 Opposite color channel sampling scheme

RGB image. From this figure, we can notice that two combinations are possible: OC-LBP_{RG} (left) and OC-LBP_{RB} (right). In OC-LBP_{RG}, the gray circle in the red channel is the central point, while the green circles in the green channel correspond to “0” sampling points and the white circles correspond to “1” sampling points, respectively. Similarly, in OC-LBP_{RB}, the blue circles correspond to “0” sampling points and the white circles correspond to “1” sampling points, respectively.

After computing the OC-LBP operator for all pixels of a given image, a total of six texture maps are generated. As depicted in Fig. 4, three intra-maps and three inter-maps are generated for each color space. Although all possible combinations of the opposite color channels allow six distinct maps, we observed that the symmetric opposing pairs are very redundant (e.g., OC-LBP_{RG} is equivalent to OC-LBP_{GR}, OC-LBP_{HS} is equivalent to OC-LBP_{SH}, and so on). Due to this redundancy, only the three more descriptive inter-maps are used.

Opponent color local salient pattern

Although OC-LBP increases the discriminability of LBP by incorporating color-texture information, it does not necessarily mimics the human visual system (HVS) behavior. To generate general-purpose descriptors that incorporates visual attention, we modify the OC-LBP by incorporating the VS information. The modified descriptor is named opponent color local salient pattern (OC-LSP). Basically, we compute the OC-LBP for all pixels of an image, obtaining the intra- and inter-channel maps of the image (see Fig. 4). In other words, being $\mathcal{L} \in \{LBP_X, LBP_Y, LBP_Z, OC-LBP_{XY}, OC-LBP_{XZ}, OC-LBP_{YZ}\}$, where XYZ represents any color space (i.e., HSV, CIE Lab, RGB, and YCbCr) normalized in the range [0, 255]. Each label $\mathcal{L}(x, y)$ corresponds to the local texture associated to the pixel $I(x, y)$. We use a VSM to generate a saliency map \mathcal{W} , where each pixel $\mathcal{W}(x, y)$ corresponds to the saliency of pixel $I(x, y)$. Figure 5a and b depicts an image and its corresponding saliency map, respectively.

The saliency map \mathcal{W} is used to weight each pixel of the map \mathcal{L} . This weighting process is used to generate a feature vector based on the histogram of \mathcal{L} weighted by \mathcal{W} . The histogram is given by the following expression:

$$\mathcal{H} = \{h_0, h_1, h_2, \dots, h_{p+1}\}, \tag{8}$$

where h_ϕ is the count of the label $\mathcal{L}(x, y)$ weighted by \mathcal{W} , as given by:

$$h_\phi = \sum_{x,y} \mathcal{W}(x, y) \cdot \delta(\mathcal{L}(x, y), \phi), \tag{9}$$

where

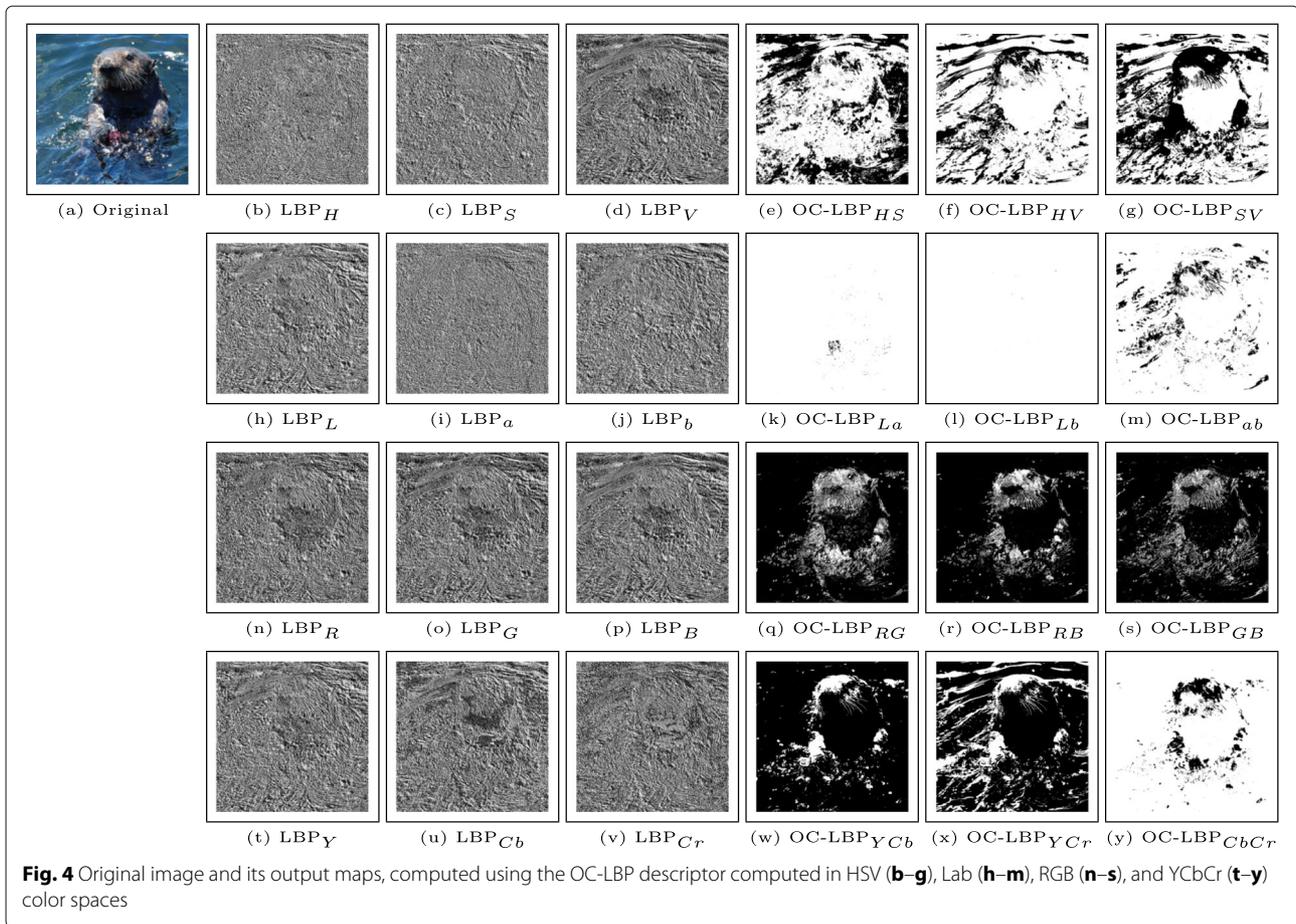
$$\delta(v, u) = \begin{cases} 1, & \text{if } v = u, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

The number of bins of \mathcal{H} is the number of distinct labels of \mathcal{L} . Therefore, we can remap each $\mathcal{L}(x, y)$ to its weighted form, generating the map $\mathcal{S}(x, y)$ that is the local salient pattern (LSP) map. Figure 5 depicts \mathcal{S} .

Opposite color local variance pattern

The use of the LBP operator (or of its variants) in IQA is based on the assumption that visual distortions affect image textures and their statistics. Particularly, images with similar distortions, at similar strengths, have textures that share analogous statistical properties. Recently, Freitas et al. [52] used a second assumption, which considers the changes in the spread of the local texture energy that are commonly observed in impaired images. For instance, a Gaussian blur impairment decreases the local texture energy, while a white noise impairment increases it. Therefore, we can use techniques that measure texture energy in RIQA algorithms.

To take into consideration the spread of the texture local energy, Freitas et al. proposed the local variance pattern (LVP) descriptor [52] for quality assessment tasks.



The LVP descriptor computes the local texture-energy according to the following formula:

$$LVP_{R,P}^{\mu}(I_c) = \left\lceil \frac{P \cdot V_{R,P}(I_c) - [LBP_{R,P}^{\mu}(I_c)]^2}{P^2} \right\rceil, \quad (11)$$

where:

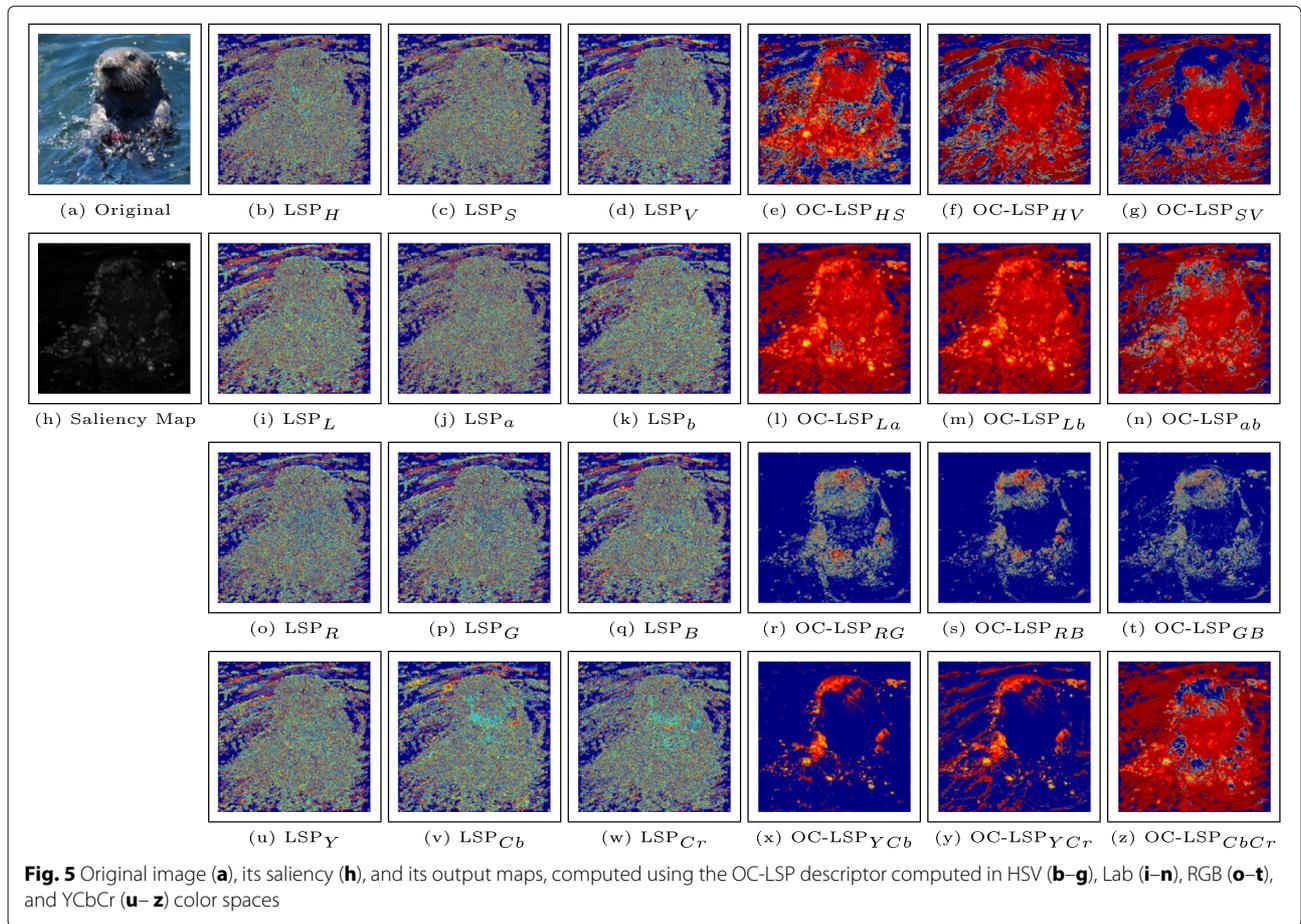
$$V_{R,P}(I_c) = \sum_{p=0}^{P-1} [f(t_p, t_c, p)]^2, \quad (12)$$

and $\lceil \cdot \rceil$ represents the operation of rounding to the nearest integer.

Figure 2 depicts the steps to extract the texture-energy information using the LVP operator. Similar to LBP operator, a LVP map is generated after computing the LVP descriptor for all pixels of a given image. A comparison between LVP and LBP maps is depicted in Fig. 6. In this figure, the first column corresponds to the reference (undistorted) image, while the three other columns correspond to images impaired with blur, white noise, and JPEG-2K distortions. The first row shows the colored images, while the second and third rows show the corresponding LBP and LVP maps, respectively. Notice

that textures are affected differently by different impairments. For instance, the LBP maps (second line of Fig. 6) corresponding to noise, blurry, and JPEG-2K compressed images have clear differences among themselves. However, the LBP map corresponding to the noise and reference images are similar. This similarity affects the discrimination between unimpaired and impaired images, affecting the quality prediction. On the other hand, the LVP channels (third line of Fig. 6) clearly show the differences between impaired and reference images.

Although the LVP descriptor presents higher discriminability (when compared with LBP), it does not incorporate color information. To take advantage of the LVP properties and include color information, we combine the OC-LBP and LVP descriptors to produce a new descriptor: the opposite color local variant pattern (OC-LVP). OC-LVP uses a sampling strategy that is similar to the strategy used by the OC-LBP descriptor (see Fig. 3), with a difference that it replaces Eq. 5 by Eq. 11. Similar to OC-LBP, OC-LVP generates six maps. As depicted in Fig. 7, three LVP intra-channel maps are generated by computing LVP independently for each color channel. Likewise, three OC-LVP inter-channels are computed by sampling



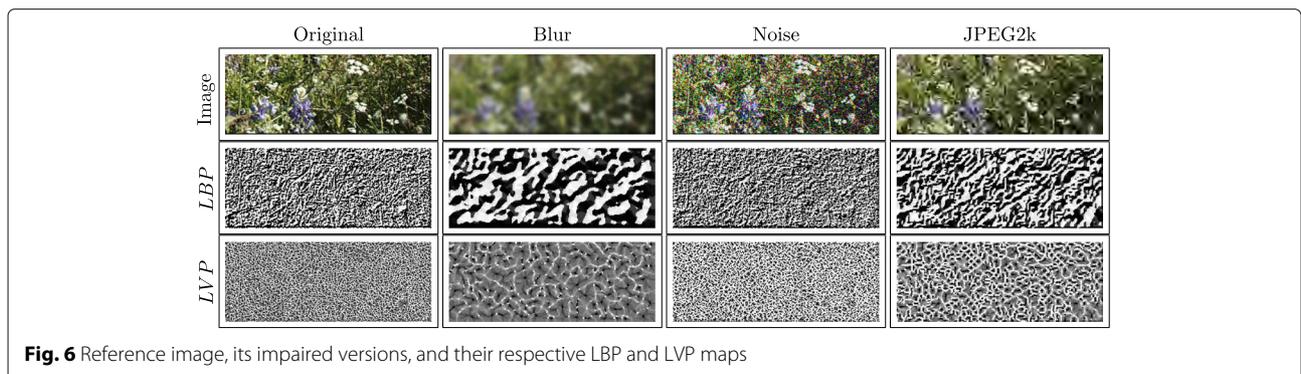
the central point in a channel and the neighboring points in another channel (across channels).

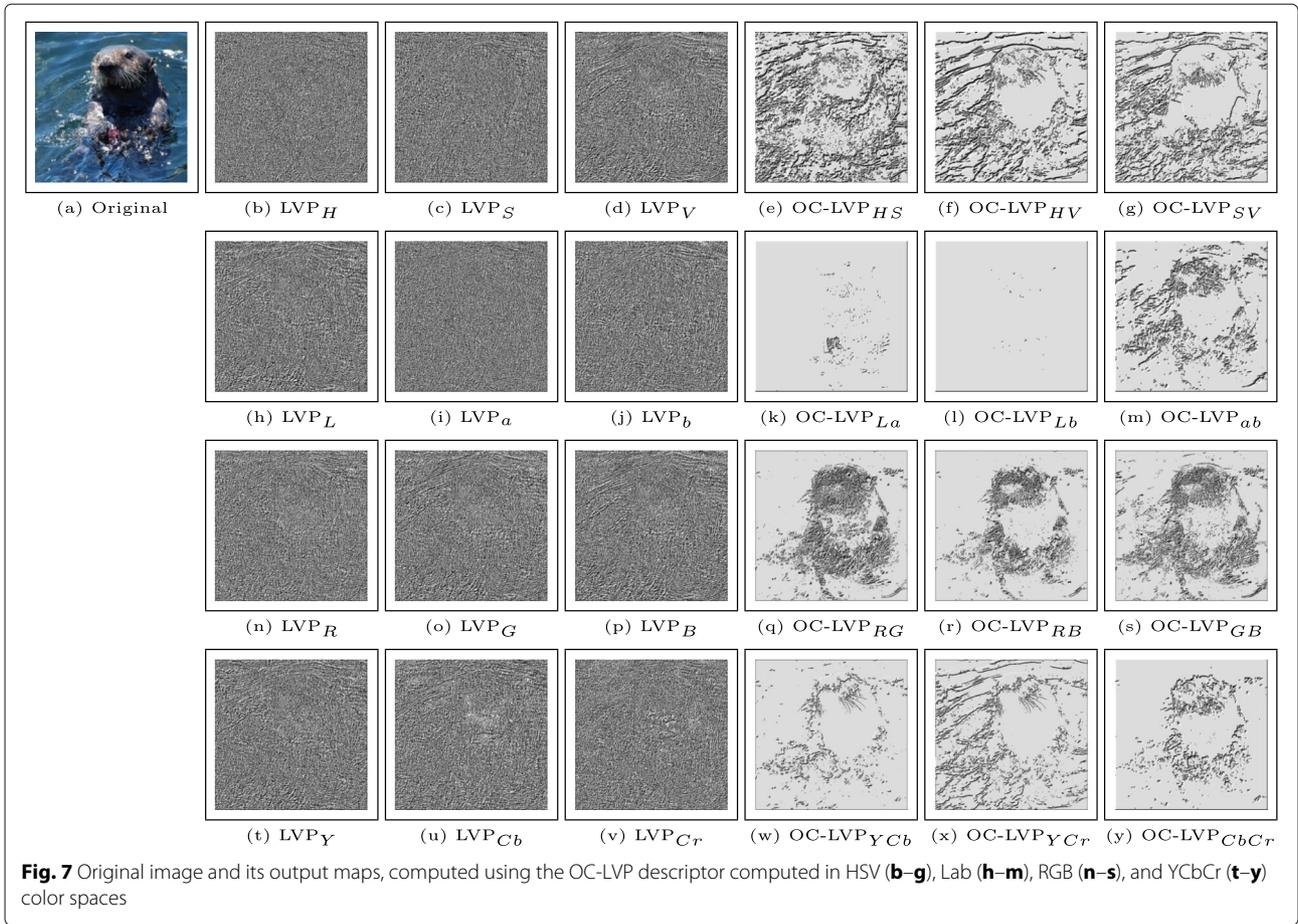
Feature extraction

The proposed RIQA method uses a supervised ML approach. The set of features is extracted, as depicted in Fig. 8. The first step of the feature extraction process consists of splitting the color channels. Using the individual color channels, we compute the OC-LSP maps. In Fig. 5, we observe that, independent of the color space,

the intra-channel maps are very similar. This similarity and the invariance between color spaces indicate that intra-channel statistics do not depend on the chosen color space.

The inter-channel maps, on the other hand, are not similar to each other. Moreover, they show considerable differences for the different color spaces. This indicates that different OC-LSP are able to extract different information, depending of the color space. Therefore, based on these observations, we use Eq. 8 to compute the histograms



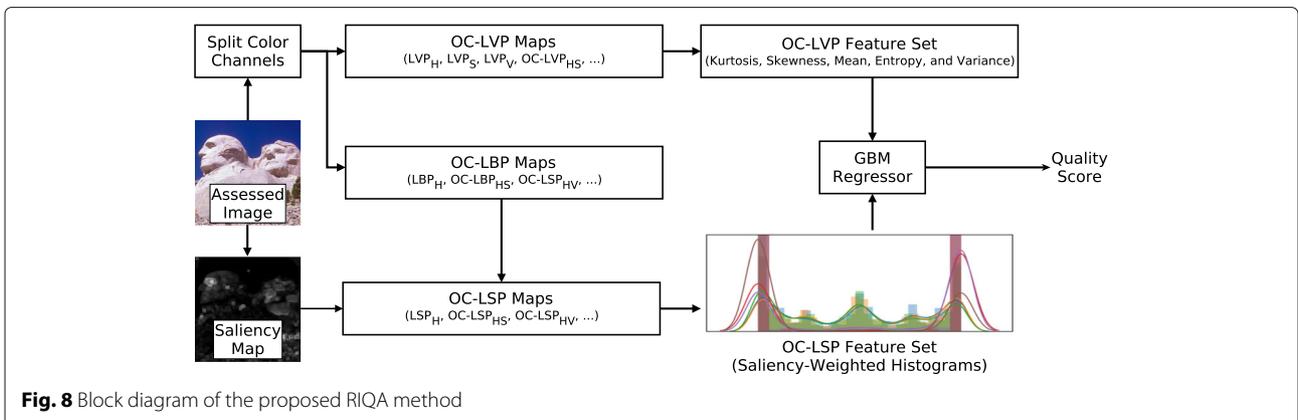


\mathcal{H} of LSP_H , $OC-LSP_{HS}$, $OC-LSP_{HV}$, $OC-LSP_{SV}$, $OC-LSP_{La}$, $OC-LSP_{Lb}$, $OC-LSP_{ab}$, $OC-LSP_{RG}$, $OC-LSP_{RB}$, $OC-LSP_{GB}$, $OC-LSP_{YCb}$, $OC-LSP_{YCr}$, and $OC-LSP_{CbCr}$ maps. The concatenation of these histograms generates the OC-LSP feature set.

Finally, the OC-LVP feature set is generated by computing the mean, variance, skewness, kurtosis, and entropy of each map, as depicted in Fig. 8. The concatenation of OC-LVP and OC-LSP feature sets generates the feature vector \vec{x} , which is used as input to a regression algorithm.

Gradient boosting machine for regression

After concatenating the OC-LVP and OC-LSP feature sets to generate the feature vector \vec{x} , we use it to predict image quality. The prediction is computed using \vec{x} as input to a gradient boosting machine (GBM). GBMs are a group of powerful ML techniques that have shown substantial success in a wide range of practical applications [53, 54]. In our application, we use a GBM regression model to map \vec{x} to the database subjective scores.



Results and discussion

In this section, we analyze the proposed method by comparing it with some of the state-of-the-art methods. Specifically, this section describes the experimental setup and configurations used in the analysis of the impact of the color space on the performance of the proposed method and in the comparisons between the proposed method and available state-of-the-art methods.

Experimental setup

There are a number of existing benchmark image quality databases. In this work, we use the following databases:

- Laboratory for Image and Video Engineering (LIVE) Image Database version 2 [60]: The database presents 982 test images, including 29 originals and 5 categories of distortions. These images are in uncompressed BMP format at several dimensions, including 480×720 , 610×488 , 618×453 , 627×482 , 632×505 , 634×438 , 634×505 , 640×512 , and 768×512 . The distortions include JPEG, JPEG 2000 (JPEG2k), white noise (WN), Gaussian blur (GB), and fast fading (FF).
- Computational and Subjective Image Quality (CSIQ) Database [28]: The database contains 30 reference images, obtained from public-domain sources, and 6 categories of distortions. These images are in $512 \times 512 \times 24$ compressed bitmap (BMP) format (PNG image data). The distortions include JPEG, JPEG 2000 (JPEG2k), white noise (WN), Gaussian blur (GB), global contrast decrements (CD), and additive Gaussian pink noise (PN). In total, there are 866 distorted images.
- Tampere Image Database 2013 (TID2013) [61]: The database has 25 reference images and 3,000 distorted images (25 reference images \times 24 types of distortions \times 5 levels of distortions). These images are in $512 \times 384 \times 24$ uncompressed BMP format. The distortions include additive Gaussian noise (AGN), additive noise in color components (ANCC), spatially correlated noise (SCN), masked noise (MN), high frequency noise (HFN), impulse noise (IN), quantization noise (QN), Gaussian blur (GB), image denoising (ID), JPEG, JPEG2k, JPEG transmission errors (JPEG+TE), JPEG2k transmission errors (JPEG2k+TE), non eccentricity pattern noise (NEPN), local block-wise distortions (LBD), intensity shift (IS), contrast change (CC), change of color saturation (CCS), multiplicative Gaussian noise (MGN), comfort noise (CN), lossy compression (LC), image color quantization with dither (ICQ), chromatic aberration (CA), and sparse sampling and reconstruction (SSR).

The Boolean Map Saliency (BMS) is used as the VSM algorithm [62]. We compare the proposed method with

a set of publicly available methods. The chosen state-of-the-art RIQA methods are the following: Codebook Representation for No-Reference Image Assessment (CORNIA) [23], Curvlet-based Quality Assessment (CQA) [50], Spatial and Spectral Entropies Quality Assessment (SSEQ) [18], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [16], local ternary patterns (LTP) [20], and No-Reference Free Energy Principle Metric (NFERM) [33]. Additionally, we also compared the proposed algorithm with three well-established FR-IQA metrics, namely PSNR, structural similarity (SSIM) [32], and reduced-reference image quality metric for contrast change (RIQMC) [49].

The trained-based RIQA methods are performed using the same training-and-testing protocol. The protocol consists of splitting each single database into two content-independent subsets (i.e., one subset for training and another for testing). To avoid overtraining and, therefore, failing to predict quality for other contents, scenes in the testing subset are not present in the training subset, and vice-versa. Considering this constraint, 20% of images are randomly selected for testing and the remaining 80% are used for training. This 80-20 split, training, and testing procedure is a simulation. We performed each simulation 1000 times, and the mean correlation value is reported. To compare the predicted and subjective quality scores, three correlation metrics were used: Spearman rank order correlation coefficient (SROCC), Pearson linear correlation coefficient (LCC), and Kendall rank order correlation coefficient (KRCC).

It is worth pointing out that each simulation is performed using all distortions in training. When the prediction performance “per distortion” is reported, the predicted data for each distortion is generated using the trained data using all distortions for training. For the training-based methods based on the support vector regression (SVR) algorithm, the training and predicting steps are implemented using the Sklearn library [63]. The SVR metaparameters are found using exhaustive grid search methods provided by Sklearn’s API. The proposed method, on the other hand, uses GBM regression implemented with the XGBoost [64] library.

Impact of color space on prediction performance

To investigate the most suitable color space for the proposed method, we perform simulations with the LIVE2 database using the HSV, Lab, RGB, and YCbCr color spaces. For comparison purposes, we also tested the algorithm using the features obtained by combining all color spaces. Table 1 shows the average LCC, SROCC, and KRCC correlation scores (CS) for 1000 simulations.

From these results, we notice that the YCbCr color space provides a statistically superior performance for almost

Table 1 Average LCC, SROCC, and KRCC of 1,000 runs of simulations, using different color spaces on LIVE, CSIQ, and TID2013 databases (Continued)

Distortion	HSV			LAB			RGB			YCbCr			ALL		
	SROCC	LCC	KRCC												
MN	0.7582	0.7388	0.5576	0.6577	0.6889	0.4800	0.5114	0.5618	0.3533	0.5920	0.6151	0.4274	0.7738	0.7917	0.5776
MS	0.0869	0.0915	0.0600	0.2115	0.1868	0.1467	0.1260	0.1092	0.0902	0.1138	0.0865	0.0835	0.1254	0.0775	0.0867
MGN	0.7844	0.7715	0.6165	0.7426	0.7453	0.5442	0.6792	0.6853	0.4908	0.7662	0.7617	0.5667	0.8769	0.8836	0.7045
NEPN	0.1929	0.1835	0.1353	0.1408	0.1683	0.0968	0.2058	0.2008	0.1436	0.1960	0.1764	0.1369	0.1985	0.2100	0.1369
QN	0.8941	0.8958	0.7267	0.8562	0.8530	0.6800	0.8215	0.8093	0.6400	0.8750	0.8460	0.7045	0.8662	0.8700	0.6912
SSR	0.8992	0.9237	0.7400	0.8946	0.9155	0.7267	0.8900	0.9266	0.7200	0.8858	0.9229	0.7200	0.9146	0.9415	0.7600
SCN	0.7262	0.7269	0.5533	0.8708	0.8804	0.7000	0.7169	0.7323	0.5467	0.8323	0.8424	0.6400	0.9023	0.9098	0.7133
ALL	0.7008	0.7553	0.5230	0.7097	0.7560	0.5214	0.6418	0.7222	0.4656	0.6888	0.7537	0.5058	0.7231	0.7746	0.5444
Average	0.7486	0.7586	0.5973	0.7407	0.7495	0.5845	0.6841	0.7051	0.5322	0.7323	0.7529	0.5837	0.7767	0.7885	0.6282

Table 2 Mean SROCC of the PSNR, SSIM, RIQMC, BRISQUE, CORNIA, CQA, SSEQ, LTP, NFERM, and the proposed metrics, obtained for 1000 simulation runs on the LIVE, CSIQ, and TID2013 databases

Database	Distortion	PSNR	SSIM	RIQMC	BRISQUE	CORNIA	CQA	SSEQ	LTP	NFERM	PROPOSED
LIVE	JPEG	0.8515	0.9481	0.7794	0.8641	0.9002	0.8257	0.9122	0.9395	0.9645	0.9325
	JPEG2k	0.8822	0.9438	0.5383	0.8838	0.9246	0.8366	0.9388	0.9372	0.9411	0.9497
	WN	0.9851	0.9793	0.6628	0.9750	0.9500	0.9764	0.9544	0.9646	0.9838	0.9845
	GB	0.7818	0.8889	0.8711	0.9304	0.9465	0.8377	0.9157	0.9530	0.9219	0.9641
	FF	0.8869	0.9335	0.6802	0.8469	0.9132	0.8262	0.9038	0.8758	0.8627	0.8977
	ALL	0.8013	0.8902	0.6785	0.9098	0.9386	0.8606	0.9356	0.9316	0.9405	0.9492
CSIQ	JPEG	0.9009	0.9309	0.7242	0.8525	0.8319	0.6506	0.8066	0.9292	0.9036	0.9331
	JPEG2k	0.9309	0.9251	0.5795	0.8458	0.8405	0.8214	0.7302	0.8877	0.9223	0.8871
	WN	0.9345	0.8761	0.4678	0.6931	0.6187	0.7276	0.7876	0.6454	0.9214	0.9346
	GB	0.9358	0.9089	0.8007	0.8337	0.8526	0.7486	0.7766	0.9244	0.8962	0.9197
	PN	0.9315	0.8871	0.3653	0.7740	0.5340	0.5463	0.6661	0.7828	0.6334	0.9461
	CD	0.8862	0.8128	0.9565	0.4255	0.4458	0.5383	0.4172	0.2082	0.3774	0.8097
	ALL	0.8088	0.8116	0.5066	0.7597	0.6969	0.6369	0.7007	0.8280	0.9142	0.8949
TID2013	AGC	0.8568	0.7912	0.3555	0.4166	0.2605	0.3964	0.3949	0.5963	0.7077	0.9217
	AGN	0.9337	0.6421	0.6055	0.6416	0.5689	0.6051	0.6040	0.6631	0.8567	0.8662
	CA	0.7759	0.7158	0.5726	0.7310	0.6844	0.4380	0.4366	0.6749	0.6357	0.5991
	CC	0.4608	0.3477	0.8044	0.1849	0.1400	0.2043	0.2006	0.1886	0.2148	0.7583
	CCS	0.6892	0.7641	0.0581	0.2715	0.2642	0.2461	0.2547	0.2384	0.3106	0.5765
	CN	0.8838	0.6465	0.6262	0.2176	0.3553	0.1623	0.1642	0.3880	0.1385	0.5100
	GB	0.8905	0.8196	0.7687	0.8063	0.8341	0.7019	0.7058	0.7465	0.8502	0.8655
	HFN	0.9165	0.7962	0.4267	0.7103	0.7707	0.7104	0.7061	0.7626	0.8797	0.9319
	ICQ	0.9087	0.7271	0.8691	0.7663	0.7044	0.6829	0.6834	0.7603	0.4804	0.7877
	ID	0.9457	0.8327	0.8661	0.5243	0.7227	0.6711	0.6716	0.7063	0.6405	0.8388
	IN	0.9263	0.8055	0.1222	0.6848	0.5874	0.4231	0.4272	0.6484	0.1735	0.6699
	IS	0.7647	0.7411	0.5979	0.2224	0.2403	0.2011	0.2013	0.3291	0.0407	0.8792
	JPEG	0.9252	0.8275	0.7293	0.7252	0.7815	0.6317	0.6284	0.6631	0.8711	0.6160
	JPEGTE	0.7874	0.6144	0.6009	0.3581	0.5679	0.2221	0.2195	0.2314	0.1281	0.8531
	JPEG2k	0.8934	0.7531	0.5967	0.7337	0.8089	0.7219	0.7205	0.7780	0.8068	0.3819
	JPEG2kTE	0.8581	0.7067	0.7189	0.7277	0.6113	0.6529	0.6529	0.6594	0.1686	0.1754
	LBD	0.1301	0.6213	0.2471	0.2833	0.2157	0.2382	0.2290	0.3813	0.1995	0.9000
	LC	0.9386	0.8311	0.5346	0.5726	0.6682	0.4561	0.4460	0.6533	0.6516	0.7738
	MGN	0.9085	0.7863	0.3751	0.5548	0.4393	0.4969	0.4897	0.6209	0.7159	0.1254
	MN	0.8385	0.7388	0.0438	0.2650	0.2342	0.2506	0.2575	0.4243	0.2238	0.8769
NEPN	0.6931	0.5326	0.1496	0.1821	0.2855	0.1308	0.1275	0.1256	0.0667	0.1985	
QN	0.8636	0.7428	0.8697	0.5383	0.4922	0.7242	0.7214	0.7361	0.7716	0.8662	
SCN	0.9152	0.7934	0.7811	0.7238	0.7043	0.7121	0.7064	0.7015	0.2181	0.9146	
SSR	0.9241	0.7774	0.6967	0.7101	0.8594	0.8115	0.8084	0.8457	0.7865	0.9023	
ALL	0.6869	0.5758	0.4439	0.5416	0.6006	0.4925	0.4900	0.6078	0.3971	0.7231	
Average		0.8377	0.7807	0.5808	0.6234	0.6262	0.5741	0.5893	0.6563	0.6084	0.7767

all distortions (23 out of 114 CS or 20.17%) and followed by Lab (13 out of 114 or 11.41%), HSV (10 out of 114 or 8.77%), and RGB (3 of 114 or 2.63%). However,

the combination of all color spaces (“ALL” label) provides the best prediction performance (65 out of 114 or 57.02%).

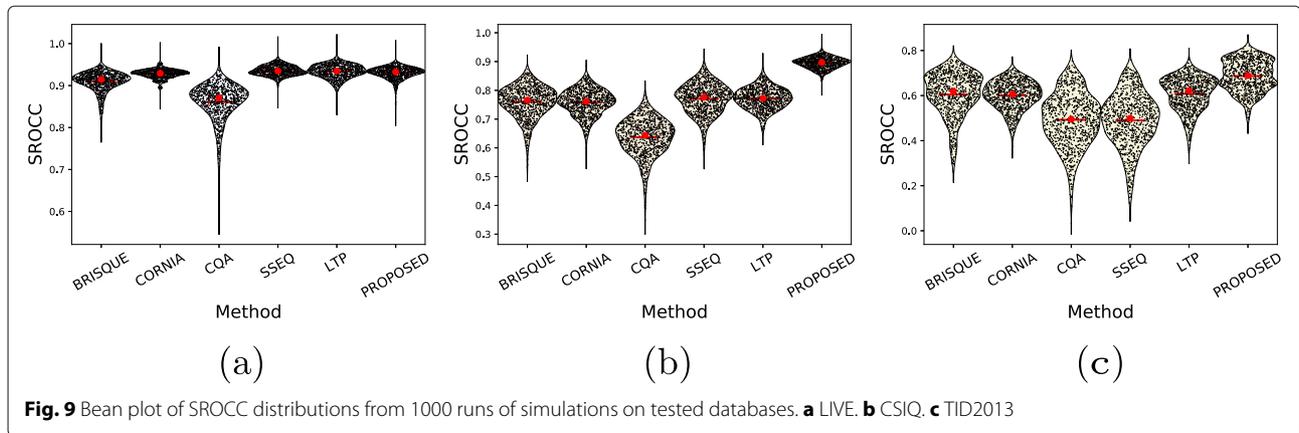


Fig. 9 Bean plot of SROCC distributions from 1000 runs of simulations on tested databases. **a** LIVE. **b** CSIQ. **c** TID2013

Prediction performance using a single database

Table 2 depicts the results for the tested methods using part of database for training and part for testing. Numbers in italics represent the best correlation values among RIQA and FR-IQA methods, while numbers in bold correspond to the best correlation values considering only the RIQA methods.

From Table 2, we can see that, for most databases, the proposed method achieves the best performance among the RIQA methods. For the LIVE2 database, the proposed method outperforms even the FR-IQA methods for JPEG2, WN, GB, and “ALL” distortions. For the PN and CD distortions of the CSIQ database, the proposed method provides a significantly better performance than the other RIQA methods. The only exception is RIQMC that obtained a mean SROCC of 0.9565, which is expected since it is a contrast-specific metric. The superior performance for PN distortions is probably due to the color-based features. The good performance for CD distortions is an important advantage of the proposed method, given that this distortion is a challenge for most RIQA methods.

For the TID2013 database, the proposed method outperforms other RIQA methods for 18 out of the 25 distortions, followed by NFERM, BRISQUE, and CORNIA. For AGC, HFN, IS, JPEG+TE, SSR, LBP, and MN distortions, the performance of the proposed method surpasses even FR-IQA methods. The performance for AGC distortions is very good, similar to what was obtained for the PN distortions of the CSIQ database. Albeit losing for RIQMC, which is a contrast-specific metric, the performance of the proposed method for CC and CD distortions of the CSIQ database is also good. This shows that the proposed method can handle contrast distortions.

Figure 9 depicts the distributions of the SROCC values computed between the subjective scores (MOS) and the predicted scores obtained using the tested RIQA methods. The bean plots of this figure are generated using the distribution of SROCC values for the set containing all database distortions (corresponding to “ALL” in Table 2).

From Fig. 9a, we notice that almost all methods (with the exception CQA) present similar distributions of SROCC scores for the LIVE database. On the other hand, SROCC values vary more for CSIQ and TID2013 databases, as can be seen in Fig. 9b, c.

Statistical difference significance test

We also conduct tests to determine the statistical significance of the differences of the coefficient values reported in Table 3. We used the Welch’s *t* test on the SROCC values corresponding to each color space, considering all distortions (“ALL” label), with a 95% confidence level. The cells in Table 3 indicate whether the value of the corresponding row is statistically better (↑), statistically inferior (↓), or statistically equivalent (⊙) to the value of the corresponding column. These results show that the proposed method has a statistically superior performance in all cases.

Performance for a cross-database validation

To investigate the generalization capability of the proposed method, we performed a cross-database validation. This validation consists of training the proposed RIQA method using all images of one database and testing them

Table 3 Welch’s *t* test performed between SROCC average values for TID2013 database: “↑” indicates that the method depicted in the row is statistically superior to the one in the column, “↓” indicates that the row is worse than the column, and “⊙” indicates that two methods have statistically the same performance

	BRISQUE	CORNIA	CQA	SSEQ	LTP	PROPOSED
BRISQUE	-	⊙	↑	↑	⊙	↓
CORNIA		-	↑	↑	↓	↓
CQA			-	⊙	↓	↓
SSEQ				-	↓	↓
LTP					-	↓
Proposed						-

Table 4 SROCC cross-database validation, when models are trained on LIVE2 and tested on CSIQ and TID2013

	DIST	BRISQUE	CORNIA	CQA	SSEQ	LTP	Proposed
CSIQ	JPEG	0.8209	0.7062	0.7129	0.8141	0.8784	0.9876
	JPEG2k	0.8279	0.8459	0.6957	0.7862	0.8914	0.9881
	WN	0.6951	0.8627	0.6596	0.4613	0.7739	0.9962
	GB	0.8311	0.8815	0.7648	0.7758	0.8712	0.9934
	ALL	0.8022	0.7542	0.7114	0.7403	0.8628	0.9914
TID 2013	JPEG	0.8058	0.7423	0.8071	0.7823	0.8472	0.8853
	JPEG2k	0.8224	0.8837	0.7724	0.8258	0.9046	0.9481
	WN	0.8621	0.7403	0.8692	0.6959	0.6881	0.9077
	GB	0.8245	0.8133	0.8214	0.8624	0.8693	0.8693
	ALL	0.7965	0.7599	0.8214	0.7955	0.8137	0.8923

on the other databases. Table 4 depicts the SROCC values obtained using LIVE as the training database and TID2013 and CSIQ as the testing databases. To perform a straightforward cross-database comparison, only similar distortions were selected from each database. In other words, we select only JPEG, JPEG2k, WN, and GB distortions of CSIQ since these distortions are also present in the training databases. The PN and CD distortions were removed from the test set and, therefore, they are not listed in Table 4. Likewise, for TID2013, only JPEG, JPEG2k, WN, and GB distortions were kept. In TID2013, the HFN distortion was chosen because it is the most similar to the WN distortion.

From Table 4, we can notice that the proposed method outperforms the other RIQA methods for the cross-database validation test. Notice that the proposed method achieves the best performance for all cases, except for one. For TID, the proposed method outperforms the other methods for four out of the five distortions, while for CSIQ, it outperforms the other methods for all five distortions. Therefore, the cross-database validation test indicates that the proposed method has a better generalization capability, when compared to the tested state-of-the-art RIQA methods.

Conclusions

In this paper, we proposed a novel NDS-GP-RIQA method based on the statistics of two new texture descriptors: the OC-LSP and OC-LVP. OC-LSP descriptor extends the capabilities of the (previous) OC-LBP operator by incorporating texture, color, and saliency information. Similarly, OC-LVP fuses OC-LBP and LVP operators to incorporate texture, color, and energy information. Quality is predicted after training a regression model using a gradient boost machine. Experimental results showed that, when compared with state-of-the-art RIQA methods, the proposed method has the best performance. More specifically, when considering a wide range

of distortions, the proposed method has a clear superiority. Since the proposed method is based on simple descriptors, it can be suitable for video quality assessment. Future works include a parallel implementation of the OC-LSP and OC-LVP descriptors.

Abbreviations

AGN: Additive gaussian noise; ANCC: Additive noise in color components; BMP: Bit map; BMS: Boolean Map Saliency; BRISQUE: Blind/referenceless image spatial quality evaluator; CA: Chromatic aberration; CC: Contrast change; CCS: Change of color saturation; CD: Contrast decrements; CN: Comfort noise; CNN: Convolutional neural networks; CORNIA: Codebook Representation for No-Reference Image Assessment; CQA: Curvlet-based quality assessment; CSIQ: Computational and subjective image quality; DCT: Discrete cosine transform; DS: Distortion-specific; FF: Fast fading; FR: Full-reference; FR-IQA: Full-reference image quality assessment; GB: Gaussian blur; GBM: Gradient boosting machine; GP: General-purpose; GP-IQA: General-purpose image quality assessment; GP-RIQA: General-purpose referenceless image quality assessment; HFN: High frequency noise; HSV: Hue, saturation, and value; HVS: Human visual system; ICQ: Image color quantization; ID: Image denoising; IN: Impulse noise; IQA: Image quality assessment; IS: Intensity shift; JPEG: Joint photographic experts group; JPEG+TE: JPEG transmission errors; JPEG2k: JPEG 2000; JPEG2k+TE: JPEG2k transmission errors; KRCC: Kendall rank order correlation coefficient; LBP: Local binary pattern; LBP: Local block-wise distortions; LC: Lossy compression; LCC: Linear correlation coefficient; LIVE: Laboratory for Image and Video Engineering; LTP: Local ternary patterns; LVP: Local variance pattern; MGN: Multiplicative Gaussian noise; ML: Machine learning; MN: Masked noise; MSD: Mean squared deviation; NDS: Non-distortion-specific; NDS-GP-RIQA: Non-distortion-specific general-purpose referenceless image quality assessment; NEPN: Non eccentricity pattern noise; NFERM: No-reference free energy principle metric; NSS: Natural scene statistic; OC-LBP: Opponent local binary patterns; OC-LSP: Opposite color local salient patterns; OC-LVP: Opposite color local variance patterns; PN: Pink noise; PNG: Portable network graphics; PSNR: Peak-to-noise ratio; QN: Quantization noise; RGB: Red, green, and blue; RIQA: Referenceless image quality assessment; RIQMC: Reduced-reference image quality metric for contrast change; RR: Reduced-reference; SCN: Spatially correlated noise; SROCC: Spearman rank order correlation coefficient; SSEQ: Spatial and spectral entropies quality assessment; SSIM: Structural similarity; SSR: Sparse sampling and reconstruction; SVR: Support vector regression; VSM: Visual saliency models; WN: White noise; YCbCr: Y luma component, color blue component relative to the green component, and color red component relative to the green component

Funding

This work was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Fundação de Apoio a Pesquisa do Distrito Federal (FAP-DF), and the University of Brasília (UnB).

Authors' contributions

PGF wrote most of the text, figures, and analysis and developed the methods described in this manuscript. WYLA was responsible for implementing the compared methods to perform the experiments. MCQF is the principal investigator in this project and has guided the research and helped write and revise this manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Department of Computer Science, University of Brasília, Campus Universitário Darcy Ribeiro, Brasília - DF 70919-970 Brazil. ²Department of Electrical Engineering, University of Brasília, Campus Universitário Darcy Ribeiro, Brasília - DF 70919-970 Brazil.

Received: 4 March 2018 Accepted: 18 July 2018

Published online: 06 August 2018

References

- Seshadrinathan K, Bovik AC (2011) Automatic prediction of perceptual quality of multimedia signals—a survey. *Multimed Tools Appl* 51(1):163–186
- Chandler DM (2013) Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Process*, vol. 2013. <https://doi.org/10.1155/2013/905685>. <https://www.hindawi.com/journals/isrn/2013/905685/>
- Telecom I (2000) Recommendation 500-10: Methodology for the subjective assessment of the quality of television pictures. ITU-R Rec. BT.500
- Wang Z, Bovik AC (2009) Mean squared error: love it or leave it? A new look at signal fidelity measures. *IEEE Signal Proc Mag* 26(1):98–117
- Moghadam A, Mohammadi P, Shirani S (2015) Subjective and objective quality assessment of image: a survey. *Majles J Electr Eng* 9(1):55–83. <https://profdoc.um.ac.ir/paper-abstract-1048833.html>
- Fang Y, Ma K, Wang Z, Lin W, Fang Z, Zhai G (2015) No-reference quality assessment of contrast-distorted images based on natural scene statistics. *Signal Process Lett IEEE* 22(7):838–842
- Bahrami K, Kot AC (2014) A fast approach for no-reference image sharpness assessment based on maximum local variation. *Signal Process Lett IEEE* 21(6):751–755
- Golestaneh SA, Chandler DM (2014) No-reference quality assessment of JPEG images via a quality relevance map. *Signal Process Lett IEEE* 21(2):155–158
- Li L, Lin W, Zhu H (2014) Learning structural regularity for evaluating blocking artifacts in jpeg images. *Signal Process Lett IEEE* 21(8):918–922
- Li L, Zhou Y, Lin W, Wu J, Zhang X, Chen B (2016) No-reference quality assessment of deblocked images. *Neurocomputing* 177:572–584
- Li L, Zhu H, Yang G, Qian J (2014) Referenceless measure of blocking artifacts by Tchebichef kernel analysis. *Signal Process Lett IEEE* 21(1):122–125
- Liu L, Hua Y, Zhao Q, Huang H, Bovik AC (2016) Blind image quality assessment by relative gradient statistics and adaboosting neural network. *Signal Process Image Commun* 40:1–15
- Li Q, Lin W, Fang Y (2016) No-reference quality assessment for multiply-distorted images in gradient domain. *IEEE Signal Process Lett* 23(4):541–545. <https://doi.org/10.1109/LSP.2016.2537321>
- Saad MA, Bovik AC, Charrier C (2012) Blind image quality assessment: a natural scene statistics approach in the DCT domain. *Image Process IEEE Trans* 21(8):3339–3352
- Moorthy AK, Bovik AC (2011) Blind image quality assessment: from natural scene statistics to perceptual quality. *Image Process IEEE Trans* 20(12):3350–3364
- Mittal A, Moorthy AK, Bovik AC (2012) No-reference image quality assessment in the spatial domain. *Image Process IEEE Trans* 21(12):4695–4708
- Saad MA, Bovik AC, Charrier C (2010) A DCT statistics-based blind image quality index. *IEEE Signal Process Lett* 17(6):583–586
- Liu L, Liu B, Huang H, Bovik AC (2014) No-reference image quality assessment based on spatial and spectral entropies. *Signal Process Image Commun* 29(8):856–863
- Freitas PG, Akamine WY, Farias MC (2016) Blind image quality assessment using multiscale local binary patterns. *J Imaging Sci Technol* 60(6):60405–1
- Freitas PG, Akamine WY, Farias MC (2016) No-reference image quality assessment based on statistics of local ternary pattern. In: 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX). IEEE. pp 1–6. <https://ieeexplore.ieee.org/abstract/document/7498959/>
- Freitas PG, Akamine WY, Farias MC (2016) No-reference image quality assessment using texture information banks. In: Intelligent Systems (BRACIS), 2016 5th Brazilian Conference On. IEEE. pp 127–132. <https://ieeexplore.ieee.org/abstract/document/7839574/>
- Ye P, Doermann D (2012) No-reference image quality assessment using visual codebooks. *Image Process IEEE Trans* 21(7):3129–3138
- Ye P, Kumar J, Kang L, Doermann D (2012) Unsupervised feature learning framework for no-reference image quality assessment. In: Computer vision and pattern recognition (CVPR), 2012 IEEE Conference On. IEEE. pp 1098–1105. <https://ieeexplore.ieee.org/abstract/document/6247789/>
- Zhang M, Muramatsu C, Zhou X, Hara T, Fujita H (2015) Blind image quality assessment using the joint statistics of generalized local binary pattern. *Signal Process Lett IEEE* 22(2):207–210
- Zhang Y, Wu J, Xie X, Shi G (2016) Blind image quality assessment based on local quantized pattern. In: Pacific Rim Conference on Multimedia. Springer. pp 241–251. https://link.springer.com/chapter/10.1007/978-3-319-48896-7_24
- Wu Q, Wang Z, Li H (2015) A highly efficient method for blind image quality assessment. In: Image processing (ICIP), 2015 IEEE International Conference On. IEEE. pp 339–343. <https://ieeexplore.ieee.org/abstract/document/7350816/>
- Wu J, Lin W, Shi G (2014) Image quality assessment with degradation on spatial structure. *Signal Process Lett IEEE* 21(4):437–440
- Larson EC, Chandler DM (2010) Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron Imaging* 19(1):011006–011006
- Charrier C, Saadane A, Fernandez-Maloigne C (2017) No-reference learning-based and human visual-based image quality assessment metric. In: 19th International Conference on Image Analysis and Processing, Catania. https://link.springer.com/chapter/10.1007/978-3-319-68548-9_23
- Sheikh HR, Bovik AC (2006) Image information and visual quality. *IEEE Trans Image Process* 15(2):430–444
- Chandler DM, Hemami SS (2007) VSNR: a wavelet-based visual signal-to-noise ratio for natural images. *IEEE Trans Image Process* 16(9):2284–2298
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *Image Process IEEE Trans* 13(4):600–612
- Gu K, Zhai G, Yang X, Zhang W (2015) Using free energy principle for blind image quality assessment. *IEEE Trans Multimed* 17(1):50–63
- Zhang L, Zhang L, Mou X (2010) RFSIM: a feature based image quality assessment metric using riesz transforms. In: Image Processing (ICIP), 2010 17th IEEE International Conference On. IEEE. pp 321–324. <https://ieeexplore.ieee.org/abstract/document/5649275/>
- Kang L, Ye P, Li Y, Doermann D (2014) Convolutional neural networks for no-reference image quality assessment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp 1733–1740. <https://ieeexplore.ieee.org/abstract/document/6909620/>
- Li J, Zou L, Yan J, Deng D, Qu T, Xie G (2016) No-reference image quality assessment using Prewitt magnitude based on convolutional neural networks. *SIVIP* 10(4):609–616
- Bosse S, Maniry D, Wiegand T, Samek W (2016) A deep neural network for image quality assessment. In: Image Processing (ICIP), 2016 IEEE International Conference On. IEEE. pp 3773–3777. <https://ieeexplore.ieee.org/abstract/document/7533065/>
- Kuzovkin I, Vicente R, Petton M, Lachaux JP, Baciu M, Kahane P, Rheims S, Vidal JR, Aru J (2017) Frequency-resolved correlates of visual object recognition in human brain revealed by deep convolutional neural

- networks. bioRxiv. <https://doi.org/10.1101/133694>. <https://www.biorxiv.org/content/early/2017/05/03/133694.full.pdf>
39. Yamins DL, DiCarlo JJ (2016) Using goal-driven deep learning models to understand sensory cortex. *Nat Neurosci* 19(3):356
 40. Zhang L, Shen Y, Li H (2014) VSI: a visual saliency-induced index for perceptual image quality assessment. *IEEE Trans Image Process* 23(10):4270–4281
 41. Farias MC, Akamine WY (2012) On performance of image quality metrics enhanced with visual attention computational models. *Electron Lett* 48(11):631–633
 42. Engelke U, Kaprykowsky H, Zepernick HJ, Ndjiki-Nya P (2011) Visual attention in quality assessment. *IEEE Signal Proc Mag* 28(6):50–59
 43. Gu K, Wang S, Yang H, Lin W, Zhai G, Yang X, Zhang W (2016) Saliency-guided quality assessment of screen content images. *IEEE Trans Multimed* 18(6):1098–1110
 44. You J, Perkis A, Hannuksela MM, Gabbouj M (2009) Perceptual quality assessment based on visual attention analysis. In: *Proceedings of the 17th ACM International Conference on Multimedia*. ACM, pp 561–564. <https://doi.org/10.1145/1631272.1631356>
 45. Le Meur O, Ninassi A, Le Callet P, Barba D (2010) Overt visual attention for free-viewing and quality assessment tasks: impact of the regions of interest on a video quality metric. *Signal Process Image Commun* 25(7):547–558
 46. Le Meur O, Ninassi A, Le Callet P, Barba D (2010) Do video coding impairments disturb the visual attention deployment?. *Signal Process Image Commun* 25(8):597–609
 47. Akamine WY, Farias MC (2014) Video quality assessment using visual attention computational models. *J Electron Imaging* 23(6):061107
 48. Ortiz-Jaramillo B, Kumcu A, Philips W (2016) Evaluating color difference measures in images. In: *Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference On*. IEEE, pp 1–6. <https://ieeexplore.ieee.org/abstract/document/7498922/>
 49. Gu K, Zhai G, Lin W, Liu M (2016) The analysis of image contrast: from quality assessment to automatic enhancement. *IEEE Trans Cybern* 46(1):284–297
 50. Liu L, Dong H, Huang H, Bovik AC (2014) No-reference image quality assessment in curvelet domain. *Signal Process Image Commun* 29(4):494–505
 51. Maenpaa T, Pietikainen M, Viertola J (2002) Separating color and pattern information for color texture discrimination. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference On*. IEEE Vol. 1, pp 668–671. <https://ieeexplore.ieee.org/abstract/document/1044840/>
 52. Freitas PG, Akamine WYL, de Farias MCQ (2017) Blind image quality assessment using local variant patterns. In: *2017 Brazilian Conference on Intelligent Systems (BRACIS)*. IEEE, pp 252–257. <https://ieeexplore.ieee.org/abstract/document/8247062/>
 53. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat*:1189–1232. https://www.jstor.org/stable/2699986?seq=1#page_scan_tab_contents
 54. Natekin A, Knoll A (2013) Gradient boosting machines, a tutorial. *Front Neurobotics* 7:21
 55. Ojala T, Pietikainen M, Maenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Anal Mach Intell IEEE Trans* 24(7):971–987
 56. He DC, Wang L (1990) Texture unit, texture spectrum, and texture analysis. *Geosci Remote Sens IEEE Trans* 28(4):509–512
 57. Ojala T, Pietikainen M, Maenpää T (2000) Gray scale and rotation invariant texture classification with local binary patterns. In: *Computer Vision-ECCV 2000*. Springer, Berlin, pp 404–420
 58. Pietikainen M, Ojala T, Xu Z (2000) Rotation-invariant texture classification using feature distributions. *Pattern Recog* 33(1):43–52
 59. Jain A, Healey G (1998) A multiscale representation including opponent color features for texture recognition. *IEEE Trans Image Process* 7(1):124–128
 60. Sheikh HR, Sabir MF, Bovik AC (2006) A statistical evaluation of recent full reference image quality assessment algorithms. *Image Process IEEE Trans* 15(11):3440–3451
 61. Ponomarenko N, Jin L, Leremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F, et al. (2015) Image database TID2013: peculiarities, results and perspectives. *Signal Process Image Commun* 30:57–77
 62. Zhang J, Sclaroff S (2016) Exploiting surroundedness for saliency detection: a boolean map approach. *IEEE Trans Pattern Anal Mach Intell* 38(5):889–902
 63. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
 64. Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In: *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*. ACM, pp 785–794. <https://doi.org/10.1145/2939672.2939785>

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
