



# Blind Image Quality Assessment Based on the Use of Saliency Maps and a Multivariate Gaussian Distribution

Christophe Charrier<sup>1(✉)</sup>, Abdelhakim Saadane<sup>2</sup>,  
and Christine Fernandez-Maloigne<sup>3</sup>

<sup>1</sup> Normandie Univ., UNICAEN, ENSICAEN, CNRS, GREYC,  
14000 Caen, France

[christophe.charrier@unicaen.fr](mailto:christophe.charrier@unicaen.fr)

<sup>2</sup> Université de Nantes, XLIM, Nantes, France

[abdelhakim.saadane@univ-nantes.fr](mailto:abdelhakim.saadane@univ-nantes.fr)

<sup>3</sup> Université de Poitiers, XLIM, Poitiers, France

[christine.fernandez@univ-poitiers.fr](mailto:christine.fernandez@univ-poitiers.fr)

**Abstract.** With the widespread use of image processing technologies, objective image quality metrics are a fundamental and challenging problem. In this paper, we present a new No-Reference Image Quality Assessment (NR-IQA) algorithm based on visual attention modeling and a multivariate Gaussian distribution to predict the final quality score from the extracted features. Computational modeling of visual attention is performed to compute saliency maps at three resolution levels. At each level, distortions of the input image are extracted and weighted by the saliency maps in order to highlight degradations of visually attracting regions. The generated features are used by a probabilistic model to predict the final quality score. Experimental results demonstrate the effectiveness of the metric and show better performance when compared to well known NR-IQA algorithms.

**Keywords:** NR-IQA · Saliency maps · MVGD

## 1 Introduction

The development of image and video processing technologies and the exponential increase of new multimedia services raise the critical issue of assessing the visual quality. From several years, a number of investigations have been conducted to design robust Image Quality Assessment (IQA) metrics. Such metrics aim at predicting image quality that well correlates with Mean Opinion Scores (MOS). No-Reference IQA (NR-IQA) are interesting as they assume no knowledge of the reference image and can be embedded in practical and real-time applications. Three approaches may be used in the design of IQA algorithms.

---

This research is supported by the ANR project #ANR-16-CE39-0013.

© Springer Nature Switzerland AG 2019

E. Ricci et al. (Eds.): ICIAAP 2019, LNCS 11752, pp. 137–147, 2019.

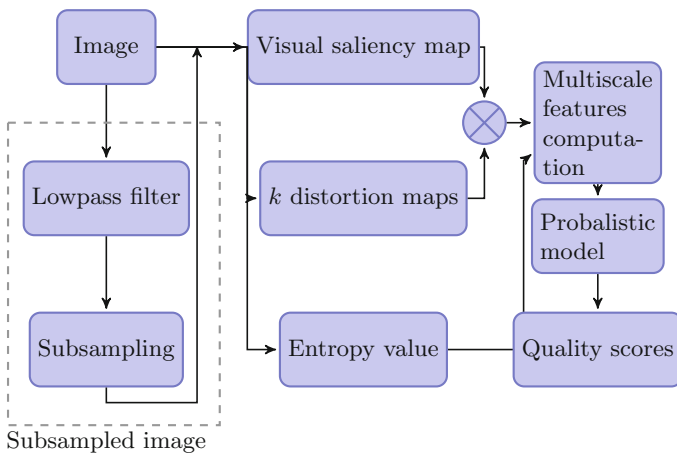
[https://doi.org/10.1007/978-3-030-30645-8\\_13](https://doi.org/10.1007/978-3-030-30645-8_13)

The first one looks to mimic the behavior of the Human Visual System (HVS). The HVS models used in this context, include relevant properties such as the contrast sensitivity function, masking effects and detection mechanisms. A number of investigations [1] have shown that these models when included in IQA algorithms, improve their performance. The second approach is well suited for assessing the quality of images distorted by known distortions. The algorithms of this approach quantify one or more distortions such as blockiness [20, 27], blur [2, 22] or ringing [9, 10] and score the image accordingly. The third and last approach is a general-purpose method. It considers that the HVS is very sensitive to structural information in the image and any loss of this structural information results in a perceptual loss of quality. To quantify loss of information, this approach uses Natural Scene Statistics (NSS). Generally, NSS-based algorithms apply a combination of learning-based approach with NSS-based extracted features. When a large ground truth is available, statistical modeling algorithms can achieve good performance. However, there is still an effort to provide to reach the subjective consistency of the HVS.

The work proposed in this paper is motivated by the interesting results of IQA when visual attention models are used. Computational visual saliency models extract regions that can attract human gaze. These regions are of a great interest in IQA. This paper presents a new NR-IQA based on the use of saliency maps to better weight the extracted distortions and combines these weighted distortions using a MultiVariate Gaussian distribution (MVGd).

## 2 The Proposed Approach

Figure 1 presents the overall synopsis of the multi-scale proposed approach, namely SABIQ (Saliency-based Blind Image Quality) index. First, a multi-scale



**Fig. 1.** Overall synopsis of the multi-scale proposed approach

decomposition is performed on the input image and a saliency map is computed at each level. The base level corresponds to the first image while the remaining ones are obtained by a low-pass filtering followed by a sub-sampling. Secondly, different distortion maps are generated at same scale levels. At each level, the Renyi entropy of the subsampled image is also computed. Thirdly, for each level, a weighting of each computed distortion map with the corresponding saliency map is performed in order to increase the strength of degradation in visually attracting areas. Finally, the combination at each level of the weighted distortion map with the computed Renyi entropy is performed to design a multiresolution distortion map. The final stage of the pipeline is a simple Bayesian model that predicts the quality score of the input image. The Bayesian approach maximizes the probability that the image has a certain quality score given the features extracted from the image. The associate posterior probability is modeled as a MultiVariate Gaussian Distribution (MVGD).

## 2.1 Visual Saliency Map

Visual attention is the ability of the HVS to rapidly direct our gaze towards regions of interest in our visual environment. Two attentional mechanisms are involved in such selection; bottom-up and top-down. Main features known to influence bottom-up attention include color, orientation, motion. Top-down attention is rather driven by the observer's experience, task and expectation. Many conducted investigations have helped in understanding visual attention and many computational saliency models have been proposed in the literature [12, 13]. A recent state-of-the-art in visual attention is given in [5]. Most of these models use bottom-up approach and are based on the Feature Integration Theory of Treisman and Gelade [28]. They compute a 2D map that highlights locations where fixations are likely to occur. These image-based (stimulus-driven) models use the same architecture but vary in the selection of characteristics used to calculate the global saliency map.

The saliency models have addressed various applications, including computer vision [21], robotics [6] and visual signal processing [7, 29]. In the context of IQA algorithms, the saliency models are intended to extract the most relevant visual features that when combined, produce a quality score highly correlated with human judgment [3].

Many research have been investigated to model the phenomenon that any human viewer can focus on attractive points at a first glance, and many saliency models have been proposed in the literature.

Saliency models can be categorized into (1) pixel-based models and (2) object-based models. The pixel-based models aim to highlight pixel locations where fixations are likely to occur. The object-based models focus on detecting salient objects in a visual scene. The majority of saliency models in the literature are pixel-based saliency models, such as ITTI [11], STB [30], PQFT [8], etc.

In this paper, the ITTI model [12] has been employed. This model combines multiscale image features into a single topographical saliency map. Three channels (Intensity, Color and Orientation) are used as low level features. First,

feature maps are calculated for each channel via center-surround differences operation. Three kinds of conspicuity maps are then obtained by across-scale combination. The final saliency map is built through combining all of the conspicuity maps.

## 2.2 Distortion Maps

Many studies have shown that image quality degradations are well measured by features of local structure [31], contrast [31, 32], multi-scale and multi-orientation decomposition [34].

**Contrast Distortion Map.** The image gradient is an interesting descriptor to capture both local image structure and local contrast [33]. Also according to this study, the partial derivatives and gradient magnitudes change with the strength of applied distortions.

Following this strategy and in order to generate the contrast distortion map, we compute both horizontal and vertical gradient component images  $\partial I/\partial x$  and  $\partial I/\partial y$  from the image  $I$ . From those two gradient images, the gradient magnitude image is computed as  $\sqrt{(\partial I/\partial x)^2 + (\partial I/\partial y)^2}$  and then modelled by a Weibull distribution. This distribution fits well the gradient magnitude of natural images [25] and its 2 parameters (the scale parameter and the shape parameter) roughly approximate the local contrast and the texture activity in the gradient magnitude map, respectively. Larger values of the scale parameter imply greater local contrast.

Yet, instead of computing the contrast on the entire image, the image is first partitioned into equally sized  $n \times n$  blocks (referred to as local image patches), then the local contrast is computed for each block yielding in final to a local contrast map  $\mathcal{M}_C$ .

**Structural Distortion Map.** The structural distortion map considered here uses structural distortion features that are extracted from both spatial and frequency information. To extract image structure information from frequency domain, the image is partitioned into equally sized  $n \times n$  local image patches and then a 2D-DCT (Discrete Cosine Transform) is applied on each patch. The feature extraction is thus locally performed in the spatio-frequency domain according to local spatial visual processing property of the HVS [4]. To capture degradation depending on directional information in the image, block DCT coefficients are modeled along three orientations (0, 45 and 90°). For each orientation, a Generalized Gaussian is fitted to the associated coefficients, and the coefficient  $\zeta$  is computed from the histogram model as  $\zeta = \sigma(X)/\mu(X)$  where  $\sigma(X)$  and  $\mu(X)$  are the standard deviation and the mean of the DCT coefficient magnitudes, respectively. In order to select the most significant map from the three generated distortion maps, the variance of  $\zeta$  is then computed for each orientation. The distortion map associated to the highest value of the variance of  $\zeta$  is finally chosen and serve as structural distortion map, namely  $\mathcal{M}_S$ .

Since the DC (Direct Coefficient) does not convey any structural information, it is removed from all computations.

**Multi-orientation Image Property Map.** It is widely admitted that the HVS is sensitive to spatial frequency and orientation. In order to capture this sensitivity, the steerable pyramid transform [26] is used.

Let  $a(i, j, f, \theta)$  be an original coefficient issued from the decomposition process located at the position  $(i, j)$  in the frequency band  $f$  and orientation band  $\theta$ . The associated squared and normalized coefficient  $r(i, j, f, \theta)$  is defined as:

$$r(i, j, f, \theta) = k \frac{a(i, j, f, \theta)^2}{\sum_{\phi \in [0, 45, 90, 135]} a(i, j, f, \phi)^2 + \sigma^2} \quad (1)$$

In this paper, four orientation bands with bandwidths of  $45^\circ$  0, 45, 90, 135 plus one isotropic lowpass filter are used yielding in five response maps  $\{R_\theta, R_{\text{iso}}, \theta \in [0, 45, 90, 135]\}$ . The distortion map associated to the highest value of the variance is finally selected and will serve as frequency variation distortion map, namely  $\mathcal{M}_F$ .

From the four orientation bands, we compute the energy ratio in order to take account the modification of local spectral signatures of an image. This approach is inspired from the quality BLIINDS2 index [24]. Each map associated to  $\theta \in [0, 45, 90, 135]$  is decomposed into equally sized  $n \times n$  blocks. For each obtained patch, the average energy in frequency band  $\theta$  models the variance corresponding to band  $\theta$  as  $e_\theta = \sigma_\theta^2$ .

For each  $\theta \in [45, 90, 135]$ , the relative distribution of energies in lower and higher bands is then computed as:

$$E_\theta = \frac{|e_\theta - 1/n \sum_{t < \theta} e_t|}{|e_\theta + 1/n \sum_{t < \theta} e_t|} \quad (2)$$

where  $1/n \sum_{t < \theta} e_t$  represents the average energy up to frequency band  $\theta$ . Three distortion maps are then generated.

The distortion map associated to the highest value of the variance of  $E_\theta$  is finally selected and serves as energy ratio distortion map, namely  $\mathcal{M}_E$ .

### 2.3 Multiscale Features Computation

In this block, each distortion map is combined with the saliency map in order to obtain a saliency-based distortion map. From each saliency-based distortion map, a pooling strategy is applied by averaging over the highest 10th percentile coefficients across the distortion map. This pooling strategy is motivated by the fact that the “worst” distortions in an image heavily influence subjective impressions and that they are concentrated in few coefficients having higher values [18]. All the obtained values are referred to as  $df^{10}(\cdot)$ , where  $(\cdot)$  represents one of the computed distortion maps  $\{\mathcal{M}_C, \mathcal{M}_S, \mathcal{M}_F, \mathcal{M}_E\}$ . In order to get information about the distribution of the distortions (over space or isolated

distortions), the 100th percentile average of the local scores is also computed. The obtained values are referred to as  $df^{100}(\cdot)$ , where  $(\cdot)$  represents one of the computed distortion maps  $\{\mathcal{M}_C, \mathcal{M}_S, \mathcal{M}_F, \mathcal{M}_E\}$ . The whole computation leads, in total, to 8 distortion features  $\{df^{10}(k), df^{100}(k)\}, \forall k \in \{\mathcal{M}_C, \mathcal{M}_S, \mathcal{M}_F, \mathcal{M}_E\}$ .

The final feature is computed at each scale level  $l$  as

$$\text{final-feature}_l^p(k) = df_l^p(k) * \text{entropy}_l \quad (3)$$

where  $p \in \{10, 100\}$ ,  $k \in \{\mathcal{M}_C, \mathcal{M}_S, \mathcal{M}_F, \mathcal{M}_E\}$ ,  $df_l^p(k)$  represents the value of the distortion value  $df^p(k)$  at level  $l$ , and  $\text{entropy}_l$  is the Renyi entropy of the associated saliency-based distortion map. This strategy yields us to include information about the anisotropy property of distortion maps. In this paper, the number of scales  $l$  is set to 3 as this value achieves the best performance.

## 2.4 Probabilistic Model and Quality Score Prediction

The computed features and the DMOS (Difference of Mean Opinion Scores) values of training images are then used by the learning block to fit a MVGD. The resulting model SABIQ is given by:

$$\text{SABIQ}(x) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \beta)^T \Sigma^{-1} (x - \beta)\right) \quad (4)$$

where  $x = (\{\text{final-feature}_l^p(k)\}, \text{DMOS})$  corresponds to the extracted features (Eq. 3) to which is added DMOS.  $\beta$  and  $\Sigma$  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 SABIQ to assess quality of image under test.

## 3 Performance Evaluation

### 3.1 Apparatus

To provide comparison of NR-IQA algorithms, two publicly available databases are used: (1) TID2013 database [23] and (2) CSIQ database [15]. Since LIVE database [14] 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. To train our model, we used LIVE database running multiple train-test sequences. For each sequence, the image database is divided into distinct training and test sets. In each train-test sequence, 80% of the LIVE IQA Database content was chosen to design the training set, and the remaining 20% were dedicated to the test set. This means each training set contains 23 reference images and their associated distorted images. The quality scores are computed using a bootstrap process with 999 replicates.

**Table 1.** SROCC values of NR-IQA models on each distortion types for the TID2013 database.

TID2013 subset		BRISQUE	BLIINDS2	DIIVINE	SSEQ	ILNIQE	CORNIA	SABIQ
Single distortion	Additive Gaussian noise	0.852	0.722	0.855	0.807	0.876	0.756	<b>0.881</b>
	Add noise in color components	0.709	0.649	0.712	0.681	0.815	0.749	<b>0.817</b>
	Spatially correlated noise	0.491	0.767	0.463	0.635	0.923	0.727	<b>0.933</b>
	Masked noise	0.575	0.512	0.675	0.565	0.512	<b>0.726</b>	<b>0.726</b>
	High frequency noise	0.753	0.824	0.878	0.860	0.868	0.796	<b>0.910</b>
	Impulse noise	0.630	0.650	<b>0.806</b>	0.749	0.755	0.767	<b>0.805</b>
	Quantization noise	0.798	0.781	0.165	0.468	<b>0.873</b>	0.016	0.851
	Gaussian blur	0.813	0.855	0.834	0.858	0.814	<b>0.921</b>	0.861
	Image denoising	0.586	0.711	0.723	0.783	0.750	<b>0.832</b>	<b>0.793</b>
	JPEG compression	0.852	0.864	0.629	0.825	0.834	0.874	<b>0.902</b>
	JPEG2000 compression	0.893	0.898	0.853	0.885	0.857	<b>0.901</b>	<b>0.901</b>
	JPEG transmission errors	0.315	0.117	0.239	0.354	0.282	<b>0.686</b>	0.351
	JPEG2000 transmission errors	0.360	0.620	0.060	0.561	0.524	<b>0.678</b>	0.622
	Non eccentricity pattern noise	0.145	0.096	0.060	0.011	0.080	<b>0.286</b>	0.102
	Local block-wise distortions	<b>0.224</b>	0.209	0.093	0.016	0.135	0.218	0.220
	Mean shift	0.124	0.128	0.010	0.108	<b>0.184</b>	0.065	0.179
Contrast change	0.040	0.150	0.460	0.204	0.014	0.182	<b>0.508</b>	
Change of color saturation	0.109	0.017	0.068	0.074	0.162	0.081	<b>0.182</b>	
Multiple dist.	Multiplicative Gaussian noise	0.724	0.716	<b>0.787</b>	0.679	0.693	0.644	0.765
	Comfort noise	0.008	0.017	0.116	0.033	0.359	<b>0.534</b>	0.379
	Lossy comp. of noisy images	0.685	0.719	0.633	0.610	0.828	<b>0.862</b>	0.838
	Color quant. with dither	0.764	0.736	0.436	0.528	0.748	0.272	<b>0.810</b>
	Chromatic aberrations	0.616	0.539	0.661	0.688	0.679	<b>0.792</b>	0.762
	Sparse sampling and reconst.	0.784	0.816	0.834	0.895	0.865	0.862	<b>0.920</b>
Cumulative subsets		0.367	0.393	0.355	0.332	0.494	0.429	<b>0.567</b>

To assess the performance of SABIQ, the Spearman Rank Order Correlation Coefficient (SROCC) is computed between DMOS values and predicted scores from six state-of-the-art opinion-aware NR-IQA methods, including BRISQUE [17], BLIINDS2 [24], DIIVINE [19], CORNIA [17], ILNIQE [33] and SSEQ [16] which are all so far widely accepted in the research community.

### 3.2 Performance Evaluation

The SROCC between predicted DMOS and subjective DMOS is reported in Table 1 for the TID2013 database. From Table 1, one observes that SABIQ performs much better than the six other NR-IQA methods when the SROCC values for the whole database is considered. This significant gain in performance is likely induced by the visual attention that is used in the weighting of distortions. When single distortions are considered, SABIQ achieves performance comparable with CORNIA and performs better than the five remaining trail quality schemes. For multiple distortions, SABIQ performs better than BRISQUE, BLIINDS2, DIIVINE and SSEQ and competes very well with CORNIA and ILNIQE.

Similar results are shown in Table 2 for CSIQ Images database. SABIQ achieves better results for 4 out of 6 distortions and outperforms all trail NR-IQA algorithms when the entire database is considered. In this case, the gain in

**Table 2.** SROCC values of NR-IQA models on each distortion types for the CSIQ Images database.

CSIQ subset	BRISQUE	BLIINDS2	DIIVINE	SSEQ	ILNIQE	CORNIA	SABIQ
JP2K	0.866	0.895	0.830	0.848	0.906	0.746	<b>0.920</b>
JPEG	0.903	0.901	0.799	0.865	0.899	0.908	<b>0.941</b>
Gaussian noise	0.252	0.379	0.176	<b>0.872</b>	0.850	0.914	0.870
Add. Gaussian pink noise	<b>0.925</b>	0.801	0.866	0.046	0.874	0.420	0.895
Gaussian blur	0.903	0.891	0.871	0.873	0.858	0.917	<b>0.920</b>
Global contrast decrement	0.029	0.012	0.396	0.200	0.501	0.302	<b>0.509</b>
Cumulative subsets	0.566	0.577	0.596	0.528	0.815	0.663	<b>0.877</b>

performance is about 7% compared to ILNIQE and is at least 32% compared to other metrics.

We also trained the methods on TID2013 excluding multi-distorted subsets (MD), then tested them on the two other datasets and the remaining MD subsets of TID2013. The results are shown in Table 3. The NR-IQA methods IL-NIQE and SABIQ clearly outperform the other trial method when trained on single distortion. When considering the LIVE database, IL-NIQE and SABIQ achieve almost the same results, which is not surprising since many existing recent NR-IQA schemes reach high correlations on that database. Furthermore, SABIQ presents the highest SROCC value with CSIQ database. All these results tend to highlight a high generalization capability of the proposed approach.

**Table 3.** SROCC values when trained on TID2013, excluding multi-distortion subsets (MD)

	LIVE	CSIQ	MD TID2013 subsets
BRISQUE	0.522	0.639	0.122
BLIINDS2	0.511	0.456	0.322
DIIVINE	0.410	0.701	0.409
SSEQ	0.230	0.630	0.098
ILNIQE	0.899	0.631	0.568
CORNIA	0.399	0.656	0.412
SABIQ	0.896	0.787	0.562

## 4 Conclusion

In this paper, we investigated how the visual attention property of the HVS can be embedded in the NR-IQA algorithm design and in which way it can improve the prediction of image quality. The proposed approach, namely SABIQ, is based on the use of computational modeling of visual attention to compute the saliency map. At each of the three levels of the multiresolution scheme, distortions of



the input image are generated and weighted by the saliency maps in order to highlight degradations of visually attracting regions. The extracted features are used by a probabilistic model to predict the final quality score. The obtained results demonstrate the effectiveness of the approach.

## References

1. Babu, R.V., Perkiş, A.: An HVS-based no-reference perceptual quality assessment of JPEG coded images using neural networks. In: Proceedings of the 2005 International Conference on Image Processing, ICIP 2005, Genoa, Italy, 11–14 September 2005, pp. 433–436 (2005). <https://doi.org/10.1109/ICIP.2005.1529780>
2. Barland, R., Saadane, A.: Blind quality metric using a perceptual map for JPEG-2000 compressed images. In: International Conference on Image Processing (ICIP) (2006)
3. Ben Amor, M., Kammoun, F., Masmoudi, N.: Improved performance of quality metrics using saliency map and CSF filter for standard coding H264/AVC. *Multimed. Tools Appl.* **77**(15), 19377–19397 (2018). <https://doi.org/10.1007/s11042-017-5393-3>
4. Blake, R., Sekuler, R.: Perception, 5th edn. McGraw-Hill Higher Education, New York (2006)
5. Borji, A., Itti, L.: State-of-the-art in visual attention modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 185–207 (2013). <https://doi.org/10.1109/TPAMI.2012.89>
6. Breazeal, C., Scassellati, B.: A context-dependent attention system for a social robot. In: 16th International Joint Conference on Artificial Intelligence (IJCAI), San-Francisco, CA, USA, pp. 1146–1153 (1999)
7. Christopoulos, C., Skodras, A., Ebrahimi, T.: The jpeg2000 still image coding system: an overview. *IEEE Trans. Consum. Electron.* **46**, 1103–1127 (2000)
8. Guo, C., Zhang, L.: A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. *IEEE Trans. Image Process.* **19**(1), 185–198 (2010). <https://doi.org/10.1109/TIP.2009.2030969>
9. Hantao, L., Klomp, N., Heynderickx, I.: A no-reference metric for perceived ringing artifacts in images. *IEEE Trans. Circuits Syst. Video Technol.* **20**(4), 529–539 (2010)
10. Hu, S., Pizlo, Z., Allebach, J.P.: JPEG ringing artifact visibility evaluation. In: Proceedings of SPIE 9016, Image Quality and System Performance XI (2014)
11. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998). <http://link.aip.org/link/?JEI/19/011006/1>
12. Itti, L., Koch, C., Niebuhr, E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998)
13. Koch, C., Ullman, S.: Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiol.* **4**(4), 219–227 (1985)

14. Laboratory for Image & Video Engineering, University of Texas (Austin): LIVE Image Quality Assessment Database (2002). <http://live.ece.utexas.edu/research/Quality>
15. Larson, E.C., Chandler, D.M.: Most apparent distortion: full-reference image quality assessment and the role of strategy. *J. Electron. Imaging* **19**(1), 011006 (2010). <https://doi.org/10.1117/1.3267105>. <http://link.aip.org/link/?JEI/19/011006/1>
16. Liu, L., Liu, B., Huang, H., Bovik, A.: No-reference image quality assessment based on spatial and spectral entropies. *Signal Process.: Image Commun.* **29**(8), 856–863 (2014). <https://doi.org/10.1016/j.image.2014.06.006>
17. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **21**(12), 4695–4708 (2012)
18. Moorthy, A.K., Bovik, A.C.: Visual importance pooling for image quality assessment. *IEEE J. Sel. Topics Signal Process.* **3**(2), 193–201 (2009). <https://doi.org/10.1109/JSTSP.2009.2015374>
19. Moorthy, A.K., Bovik, A.C.: Blind image quality assessment: from natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **20**(12), 3350–3364 (2011)
20. Muijs, R., Kirenko, I.: A no-reference blocking artifact measure for adaptive video processing. In: *European Signal Processing Conference (Eusipco)* (2005)
21. Navalpakkam, V., Itti, L.: An integrated model of top-down and bottom-up attention for optimizing detection speed. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 2049–2056 (2006)
22. ParvezSazzad, Z., Kawayoke, Y., Horita, Y.: No-reference image quality assessment for JPEG-2000 based on spatial features. *Signal Process.: Image Commun.* **23**(4), 257–268 (2008)
23. Ponomarenko, N., Carli, M., Lukin, V., Astola, J., Battisti, F.: Color image database for evaluation of image quality metrics. In: *International Workshop on Multimedia Signal Processing, Australia*, pp. 403–408, October 2008
24. Saad, M., Bovik, A.C., Charrier, C.: Blind image quality assessment: a natural scene statistics approach in the DCT domain. *IEEE Trans. Image Process.* **21**(8), 3339–3352 (2012)
25. Scholte, H.S., Ghebreab, S., Waldorp, L., Smeulders, A.W.M., Lamme, V.A.F.: Brain responses strongly correlate with weibull image statistics when processing natural images. *J. Vis.* **9**(4), 29 (2009). <https://doi.org/10.1167/9.4.29>
26. Simoncelli, E.P., Freeman, W.T.: The steerable pyramid: a flexible architecture for multi-scale derivative computation. In: *Proceedings. In: International Conference on Image Processing (ICIP)*, vol. 3, pp. 444–447, October 1995. <https://doi.org/10.1109/ICIP.1995.537667>
27. Song, X., Yang, Y.: A new no-reference assessmet metric of blocking artefacts on HVS masking effect. In: *International Congress on Image and Signal Processing*, pp. 1–6 (2009)
28. Treisman, A.M., Gelade, G.: A feature-integration theory of attention. *Cogn. Psychol.* **12**, 97–136 (1980)
29. Walther, D., Rutishauser, U., Koch, C., Perona, P.: Selective visual attention enables learning and recognition of multiple objects in cluttered scenes. *Comput. Vis. Image Underst.* **100**, 42–63 (2005)
30. Walther, D., Koch, C.: Modeling attention to salient proto-objects. *Neural Netw.* **19**(9), 1395–1407 (2006). <https://doi.org/10.1016/j.neunet.2006.10.001>. <http://www.sciencedirect.com/science/article/pii/S0893608006002152>. *Brain and Attention*

31. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
32. Xue, W., Mou, X., Zhang, L., Bovik, A., Feng, X.: Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Trans. Image Process.* **23**(11), 4850–4862 (2014)
33. Zhang, L., Zhang, L., Bovik, A.: A feature-enriched completely blind image quality evaluator. *IEEE Trans. Image Process.* **24**(8), 2579–2591 (2015)
34. Zhang, Y., Chandler, D.: No-reference image quality assessment based on log-derivative statistics of natural scenes. *J. Electron. Imaging* **22**(4), 1–23 (2013)