

Saliency-Guided Local Full-Reference Image Quality Assessment

Domonkos Varga 

Ronin Institute, Montclair, NJ 07043, USA; domonkos.varga@ronininstitute.org

Abstract: Research and development of image quality assessment (IQA) algorithms have been in the focus of the computer vision and image processing community for decades. The intent of IQA methods is to estimate the perceptual quality of digital images correlating as high as possible with human judgements. Full-reference image quality assessment algorithms, which have full access to the distortion-free images, usually contain two phases: local image quality estimation and pooling. Previous works have utilized visual saliency in the final pooling stage. In addition to this, visual saliency was utilized as weights in the weighted averaging of local image quality scores, emphasizing image regions that are salient to human observers. In contrast to this common practice, visual saliency is applied in the computation of local image quality in this study, based on the observation that local image quality is determined both by local image degradation and visual saliency simultaneously. Experimental results on KADID-10k, TID2013, TID2008, and CSIQ have shown that the proposed method was able to improve the state-of-the-art's performance at low computational costs.

Keywords: full-reference image quality assessment; visual saliency; image features



Citation: Varga, D. Saliency-Guided Local Full-Reference Image Quality Assessment. *Signals* **2022**, *3*, 483–496. <https://doi.org/10.3390/signals3030028>

Academic Editors: Francisco Martínez González and Mohammed K. A. Kaabar

Received: 28 April 2022

Accepted: 29 June 2022

Published: 11 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Image quality assessment (IQA) is a hot research topic in the image processing community, since it can be applied in a wide range of practical and important applications, such as the optimization of computer vision system parameters [1] and monitoring the quality of image displays [2], it is also helpful in the process of benchmarking image and video compression and denoising algorithms [3,4]. In practice, the final perceivers and users of digital images are human beings. As a consequence, the most reliable way of evaluating the perceptual quality of digital images is subjective user studies involving a group of vision experts and human observers either in a laboratory environment [5] or a crowdsourcing experiment [6]. Further, these subjective evaluation methods result in publicly available benchmark IQA databases, which can be used in objective IQA. Namely, objective IQA tries to construct mathematical algorithms that estimate the perceptual quality of digital images as consistently as possible with human judgement. To help the development of IQA methods, publicly available benchmark databases, such as TID2013 [7] and CSIQ [8], contain digital images with their mean opinion score (MOS) values, which are the averages of individual human quality ratings. Traditionally, IQA is divided into three classes [9]: full-reference (FR), no-reference (NR), and reduced-reference (RR). The main difference is the degree of access to the reference, i.e., the distortion-free images. Namely, FR-IQA algorithms have full access to the reference image, while NR-IQA ones have no access to the them, and RR-IQA methods have limited access to the reference images.

Human vision research has increased the understanding of the human visual system. During observation of an image, people tend to focus on the visually significant part of a scene. As a consequence, the content of the image is not treated equally. This is why considering visual saliency in FR-IQA has gained a lot of attention in the literature [10–12]. In the literature, the traditional and well-known PSNR [13] was first improved using visual saliency computation by dividing the input images into different regions and assigning weights to them according to their estimated saliency. This work was followed by a line of papers utilizing visual saliency from different features. Specifically, Zhang et al. [14]

utilized the phase congruency feature [15] as a visual saliency measure to create weights for defined feature similarity metrics. Wang and Li [16] compiled local distortion maps from the reference and distorted images. To quantify visual quality, these local distortion maps were pooled together by using saliency as a weighting function. The common point of algorithms using visual saliency is that visual saliency is used as a weighting function in the comparison of local image quality maps.

1.1. Motivation and Contributions

Over the evolution of the human race, humans have evolved to pay different amounts of attention to different regions of a visual scene as it is viewed. This mechanism obviously affects human perceptual quality judgements. Previously published FR-IQA methods [14,16–18] utilize visual saliency in the final pooling stage of the algorithm. To be more specific, visual saliency is used as weight in the weighted averaging of local image quality scores, emphasizing image regions that are more salient to the human visual system. The main contribution of this paper was the following. As opposed to other previously published algorithms, visual saliency was utilized by this work in the computation of local image quality and not in the pooling stage, which is usual in the literature. In fact, the human visual system perceives image degradation with higher probability in visually significant regions and tends to neglect degradation in visually less significant regions. In other words, local quality degradation is simultaneously influenced by both local image degradation and visual saliency. The effect of visual saliency on the perception of local image quality is demonstrated below. If humans observe Figure 1a, they tend to focus their attention on the child's face rather than on the vegetation in the background. To be specific, we added the same amount of noise to two different visually significant areas of the image. In Figure 1a, it was added to the region containing the the vegetation in the left corner, while it was added to the face in Figure 1b. Comparing these two figures, the human visual system was less likely to perceive the degradation if it was not in a visually salient area.

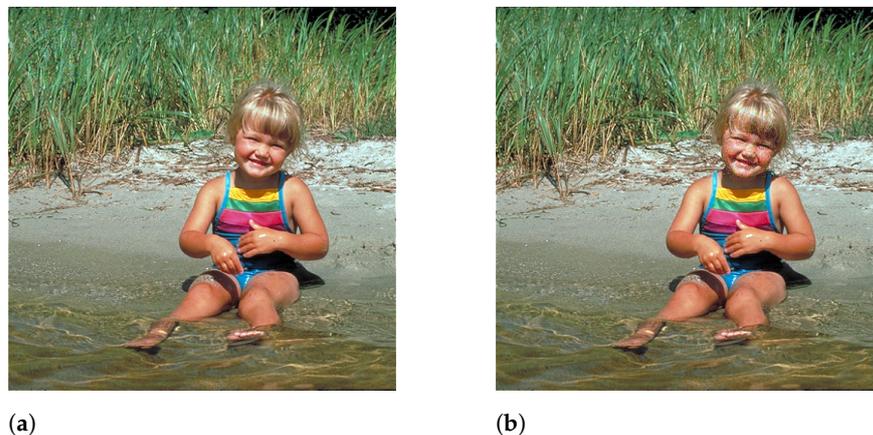


Figure 1. HVS's perception of local image quality in relation to visual saliency: (a) degradation in non-salient region, (b) degradation in salient region.

Based on the above-mentioned observations, a visual-saliency-guided FR-IQA framework was introduced where the visual saliency adaptively adjusted local image quality. Within this framework, ESSIM (edge strength similarity-based image quality metric) [19] method was further developed. Furthermore, it was demonstrated that the proposed method was able to improve state-of-the-art performance at low computational costs on four large and widely accepted IQA benchmark databases, i.e., KADID-10k [20], TID2013 [7], TID2008 [21], and CSIQ [8].

1.2. Organization of the Paper

After this introduction, related and previous papers are reviewed in Section 2. Next, our proposed method is described in Section 3. Databases, evaluation metrics, and implementation details are given in Section 4. Numerical, experimental results, and a comparison to the state-of-the-art is presented in Section 5. At the end, a conclusion is drawn in Section 6.

2. Related Work

As already mentioned, FR-IQA algorithms predict the perceptual quality of distorted images with full access to the reference images. The traditional and probably the simplest methods for FR-IQA are mean squared error (MSE) and peak signal-to-noise ratio (PSNR), which rely on the measurement of the distortion energy of the images. However, their performance lags behind other more sophisticated metrics, since their results are not consistent with human quality perception [22]. Hence, many FR-IQA methods have tried to build on the properties of the human visual system (HVS) to compile effective algorithms. For example, the structural similarity index (SSIM) [23] utilizes the observation that HVS is sensitive to contrast and structural distortions. More specifically, SSIM [23] applies local sliding windows on the reference and distorted images and measures luminance, contrast, and structure similarity using predefined functions containing average, variance, and covariance computations. Finally, the perceptual quality of the distorted image is obtained by taking the arithmetic mean of the similarity values of the local sliding windows. Brunet et al. [24] investigated the mathematical properties of SSIM and pointed out that SSIM satisfies the identity of indiscernibles and exhibits symmetry properties, excluding the triangle inequality. SSIM has become quickly popular in the signal processing community and gained a significant amount of attention for further research [25]. For example, Wang et al. [26] extended SSIM into MS-SSIM by conducting multi-scale processing. Further, Sampat et al. [27] utilized complex wavelet domains to define a new structural similarity index. In contrast, Chen et al. [28] modified SSIM to compare edge information between the reference and the distorted images. In addition to structural degradation, image gradient is also a popular feature in FR-IQA. Specifically, Liu et al. [29] defined an FR-IQA metric using gradient similarity between the reference and distorted images. Similarly, Zhu and Wang [30] utilized gradient similarity, but scale information was also incorporated into their metric.

Motivated by the success of deep learning in many image processing tasks [31–33], deep learning has also appeared in FR-IQA. For example, significant amount of proposed works compare deep activations of convolutional neural networks (CNNs) for a reference-distorted image pair to establish an FR quality metric. For example, Amirshahi et al. [34] extracted feature maps using an AlexNet [35] CNN. Subsequently, the extracted feature maps were compared using histogram intersection kernels [36] and a similarity value was produced for each feature map. To obtain perceptual quality, the arithmetic mean of these similarity values was taken. Later, this approach was developed further [37] by replacing the histogram intersection kernels by traditional image similarity metrics. Another line of work focused on training end-to-end deep architectures on large IQA benchmark databases [38–40]. Chubarau et al. [39] investigated vision transformers and self-attention for FR-IQA. Namely, a context-aware sampling procedure was introduced to extract patches from the reference and the distorted images. Next, the patches were encoded by a vision transformer.

Another class of FR-IQA methods uses existing and available FR-IQA metrics to construct a new metric with improved performance. For instance, Okarma [41] studied the characteristics of three different metrics. In addition, the authors introduced a combined metric by taking the exponentiated product of the examined three metrics. In contrast, other proposals utilized optimization techniques to find optimal weights for a linear combination of already existing FR-IQA techniques [42–44]. Instead of optimization, Lukin et al. [45]

trained a neural network from scratch using traditional FR-IQA metrics as image features to obtain improved estimation performance.

3. Proposed Method

In this section, our proposed method is introduced and described. Specifically, we first gave an overview about the ESSIM [19] method to better understand our contribution to the saliency-guided determination of local image quality. In the following, the ESSIM [19] algorithm is first briefly described, and then the saliency-guided SG-ESSIM is described in detail.

3.1. ESSIM Method

To quantify quality degradation of a distorted image (g) with respect to a reference image (f), the ESSIM is defined as

$$ESSIM(f, g) = \frac{1}{N} \sum_{i=1}^N \frac{2E(f, i)E(g, i) + C}{(E(f, i))^2 + (E(g, i))^2 + C}, \tag{1}$$

where C is a small constant to avoid the division by zero. To be more specific, C has two roles in Equation (1): as already mentioned, it is necessary to avoid having a denominator equal to zero; secondly C is also a scaling factor. If $C \rightarrow \infty$, the result of similarity measure is 1. Further, $E(f, i)$ is used to characterize the edge strength around pixel i in the reference image f , and is determined as

$$E(f, i) = \max(E_i^{1,3}(f), E_i^{2,4}(f)), \tag{2}$$

where $E_i^{2,4}(f)$ is the edge strength in the diagonal directions and is determined as

$$E_i^{2,4}(f) = |\partial f_i^2 - \partial f_i^4|^p. \tag{3}$$

Similarly, $E_i^{1,3}(f)$ stands for the edge strength but in the vertical and horizontal directions and is computed as

$$E_i^{1,3}(f) = |\partial f_i^1 - \partial f_i^3|^p. \tag{4}$$

In Equations (3) and (4), ∂f_i^j stands for the directional derivatives at pixel i in reference image f . The applied directions in the ESSIM are the following: $0^\circ, 45^\circ, 90^\circ, 135^\circ$. Further, the fractional derivatives in the ESSIM are implemented with the help of 5×5 Scharr operators [46]. Moreover, p is used to adapt the edge strength in the algorithm. The edge strength of distorted image g around pixel i is defined as

$$E(g, i) = \begin{cases} E_i^{1,3}(g), & \text{if } E(f, i) = E_i^{1,3}(f) \\ E_i^{2,4}(g), & \text{if } E(f, i) = E_i^{2,4}(f) \end{cases} \tag{5}$$

in order to guarantee that the edge strengths of f and g are compared in the same direction. Otherwise, $E_i^{1,3}(g)$ and $E_i^{2,4}(g)$ are determined similarly to Equations (3) and (4).

3.2. SG-ESSIM Method

As one can observe from the derivation of the ESSIM [19], the impact of visual saliency was not incorporated into these metrics. In this section, we describe how we enhanced the ESSIM by using visual saliency in the measurement of local image quality. Hence, it was highlighted that local image quality degradation was jointly characterized by the objective degradation of edge strength and significance.

As proposed by the authors of the ESSIM [19], our algorithm first determined the edge strength maps of the reference and distorted images as recommended in the ESSIM [19].

Subsequently, it calculated the saliency-guided local image quality map. Finally, it pooled the local image quality map to obtain the overall perceptual quality of the distorted image.

Using Equations (2) and (5), which describe $E(f, i)$ and $E(g, i)$ in the ESSIM, our visual saliency-guided local similarity map is defined as

$$SM(E(f, i), E(g, i)) = \frac{2 \cdot E(f, i) \cdot E(g, i) + H(V(i))}{(E(f, i))^2 + (E(g, i))^2 + H(V(i))} \tag{6}$$

where $V(i)$ stands for the visual saliency measure at pixel location i . Moreover, $H(\cdot)$ is a decreasing function defined as

$$H(V(i)) = K \cdot e^{-\frac{V(i)}{h}}, \tag{7}$$

where e is Euler’s number, K is a scaling factor, and h is an attenuation factor (Equation (7) is a decreasing function because the similarity metric defined by Equation (9) also needs to be regulated). To be more specific, $V(i)$ takes the edge strength maps of the reference and distorted images and returns their pixel-wise maximum:

$$V(i) = \max(E(f, i), E(g, i)). \tag{8}$$

Equation (8) implies that regions with strong edges are more salient to the human visual system than those with weak edges since edge information conveys essential information about the visual scene to humans [47]. A larger value of $V(i) = \max(E(f, i), E(g, i))$ indicated a higher visual saliency at pixel location i . Moreover, the computation of $V(i)$ did not result in a significant increase in the computational costs since $E(f, i)$ and $E(g, i)$ were determined anyway in the ESSIM [19]. Substituting Equations (7) and (8) into Equation (9), we obtain the following equation for visual saliency guided local similarity map:

$$SM(E(f, i), E(g, i)) = \frac{2 \cdot E(f, i) \cdot E(g, i) + K \cdot e^{-\frac{\max(E(f, i), E(g, i))}{h}}}{(E(f, i))^2 + (E(g, i))^2 + K \cdot e^{-\frac{\max(E(f, i), E(g, i))}{h}}}. \tag{9}$$

Finally, the overall predicted perceptual quality score of the image was estimated by averaging the visual saliency-guided (SG) local similarity map

$$SG - ESSIM = \frac{1}{N} \sum_{i=1}^N SM(E(f, i), E(g, i)), \tag{10}$$

where N is the total number of pixels.

The hyperparameters of the SG-ESSIM, which have to be set, were K and h . These parameters were determined on a subset of the TID2008 [21] database, which contained 5 random reference images and 340 corresponding distorted images. Specifically, those hyperparameters were chosen where the highest Spearman’s rank order correlation coefficient between the ground truth and the predicted values was obtained. As a result, we chose $K = 51,000$ and $h = 0.5$ in our MATLAB implementation. Since the SG-ESSIM was not a machine-learning-based approach, this random subset was also used in the evaluation process.

4. Materials

In this section, the description of the applied benchmark IQA databases are given first. Second, we give the applied evaluation metrics and protocol. Finally, the implementation details of the proposed method are given.

4.1. Databases

In this study, four large publicly available IQA benchmark databases (KADID-10k [20], TID2013 [7], TID2008 [21], and CSIQ [8]) were used to evaluate and compare the state-of-the-

art to the proposed method. The most important information about the applied databases is summarized in Table 1. Namely, these databases contained a small set of distortion-free reference images from which distorted images were derived using different distortion types at different distortion levels. Moreover, the distorted images were annotated with subjective quality ratings. Therefore, they were suitable and accepted for the evaluation and ranking of FR-IQA metrics in the literature [48]. Figure 2 depicts the empirical MOS distributions in the applied databases.

Table 1. Summary of benchmark databases used in this study.

Database	Ref. Images	Dist. Images	Dist. Types	Dist. Levels
KADID-10k [20]	81	10,125	25	5
TID2013 [7]	25	3000	24	5
TID2008 [21]	25	1700	17	4
CSIQ [8]	30	866	6	4–5

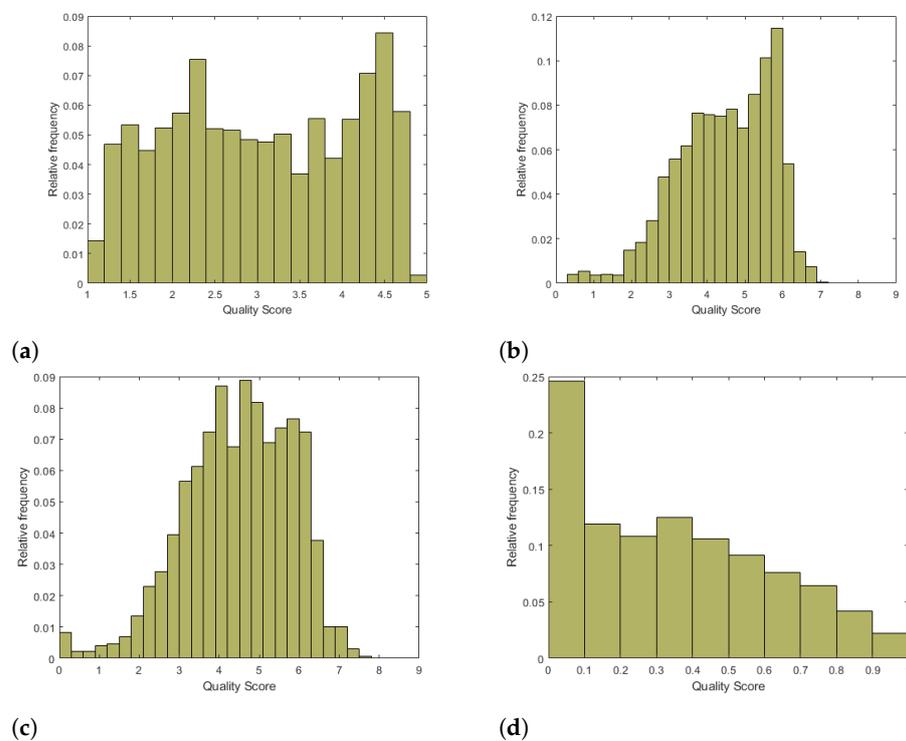


Figure 2. Empirical MOS distributions in benchmark IQA databases: (a) KADID-10k [20], (b) TID2013 [7], (c) TID2008 [21], and (d) CSIQ [8].

4.2. Evaluation Metrics and Protocol

Ranking of FR-IQA methods relies on measuring the correlation between predicted and ground truth quality scores of benchmark IQA databases, such as KADID-10k [20]. To be more specific, the performance of an FR-IQA metric was evaluated using three different criteria in this study: Pearson’s linear correlation coefficient (PLCC), Spearman’s rank order correlation coefficient (SROCC), and Kendall’s rank order correlation coefficient (KROCC). In addition, a nonlinear logistic regression was applied before the calculation of the PLCC as recommended in [49], because nonlinear relationship exists between the ground truth and predicted scores:

$$Q = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2(Q_p - \beta_3)}} \right) + \beta_4 Q_p + \beta_5, \tag{11}$$

where Q and Q_p are the fitted and predicted scores, respectively. In addition, the regression described by Equation (11) is determined by the β_i ($i \in \{1, 2, \dots, 5\}$) parameters.

4.3. Implementation Details

MATLAB R2021a was used to implement the proposed FR-IQA metric using the functions of the Image Processing Toolbox. The computer configuration applied in our experiments is outlined in Table 2.

Table 2. Computer configuration applied in our experiments.

Computer model	STRIX Z270H Gaming
Operating system	Windows 10
Memory	15 GB
CPU	Intel(R) Core(TM) i7-7700K CPU 4.20 GHz (8 cores)
GPU	Nvidia GeForce GTX 1080

5. Experimental Results and Analysis

In this section, our experimental results and analysis are presented using four common benchmark IQA databases: KADID-10k [20], TID2013 [7], TID2008 [21], and CSIQ [8]. More specifically, the proposed *SG-ESSIM* method was compared to thirteen other state-of-the-art FR-IQA metrics (2stepQA [50], CSV [51], DISTS [52], ESSIM [19], GSM [29], IW-SSIM [16], MAD [8], MS-SSIM [26], PSNR, ReSIFT [53], RVSIM [54], SSIM [23], SUMMER [55]), whose original source codes were made publicly available by their authors, on the above-mentioned databases in terms of correlations with ground truth quality ratings. Furthermore, their hyperparameter values (if any) were applied in the examined state-of-the-art methods, which were recommended as default values by the original authors in their MATLAB implementations. In addition to this, a detailed comparison of TID2013 [7] and TID2008 [21] with respect to different noise types and noise levels is also presented.

The comparison, in terms of the PLCC, SROCC, and KROCC, in KADID-10k [20], TID2013 [7], TID2008 [21], and CSIQ [8] is summarized in Tables 3 and 4. From these experimental, numerical results, it could be seen that the proposed *SG-ESSIM* was able to provide the best results in terms of the SROCC and KROCC on KADID-10k [20] and the PLCC on TID2013 [7], respectively. Moreover, it gave the best results for all correlation values in TID2008 [21]. On CSIQ [8], the second best results in terms of the SROCC and KROCC were achieved by the proposed method. To illustrate the overall performance of the examined methods on these databases, direct and weighted average (using the numbers of images as weights) performance values are summarized in Table 5. It could be seen that the proposed method was able to deliver the second best results in terms of the SROCC and KROCC, both in direct and weighted averages. In Tables 6 and 7, detailed results could be seen on TID2013 [7] and TID2008 [21] with respect to the different distortion types. Similarly, Tables 8 and 9 illustrate the detailed results with respect to the noise levels. TID2013 [7] contained images with 24 different distortion types, such as AGN (additive Gaussian noise), ANC (additive noise in color components), SCN (spatially correlated noise), MN (masked noise), HFN (high frequency noise), IN (impulse noise), QN (quantization noise), GB (Gaussian blur), DEN (image denoising), JPEG (JPEG compression), JP2K (JPEG2000 compression), JGTE (JPEG transmission errors), J2TE (JPEG2000 transmission errors), NEPN (noneccentricity pattern noise), BLOCK (local block-wise distortions of different intensity), MS (mean shift), CC (contrast change), CCS (change in color saturation), MGN (multiplicative Gaussian noise), CN (comfort noise), LCNI (lossy compression of noisy images), ICQD (image color quantization with dither), CA (chromatic aberrations), and SSR (sparse sampling and reconstruction). TID2008 [21] consisted of the first 17 distortion types from TID2013 [7]. From the presented results, it could be observed that *SG-ESSIM* was able to produce the best SROCC values on five distortion types of TID2013 [7] and on four distortion types of TID2008 [21]. On the other hand, it provided the second best results on nine distortion types of TID2013 [7] and on seven distortion types of TID2008 [21]. Moreover, the *SG-ESSIM* was able to produce the best SROCC values on three distortion levels of TID2013 [7], while it gave the second best results on the remaining two distortion levels. In TID2008 [21], it provided the best results on three distortion levels and the second best result on the remaining distortion level.

Table 3. Comparison of the *SG-ESSIM* to several other state-of-the-art algorithms on KADID-10k [20] and TID2013 [7]. The highest values are typed in **bold**, while the second highest ones are underlined.

FR-IQA Metric	KADID-10k [20]			TID2013 [7]		
	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC
2stepQA [50]	0.768	0.771	0.571	0.736	0.733	0.550
CSV [51]	0.671	0.669	0.531	<u>0.852</u>	0.848	0.657
DISTS [52]	0.809	0.814	0.626	0.759	0.711	0.524
ESSIM [19]	0.644	<u>0.823</u>	<u>0.634</u>	0.740	0.797	0.627
GSM [29]	0.780	0.780	0.588	0.789	0.787	0.593
IW-SSIM [16]	0.781	0.756	0.524	0.832	0.778	0.598
MAD [8]	0.716	0.724	0.535	0.827	0.778	0.600
MS-SSIM [26]	<u>0.819</u>	0.821	0.630	0.794	0.785	0.604
PSNR	0.479	0.676	0.488	0.616	0.646	0.467
ReSIFT [53]	0.648	0.628	0.468	0.630	0.623	0.471
RVSIM [54]	0.728	0.719	0.540	0.763	0.683	0.520
SSIM [23]	0.670	0.671	0.489	0.618	0.616	0.437
SUMMER [55]	0.719	0.723	0.540	0.623	0.622	0.472
<i>SG-ESSIM</i>	0.739	0.838	0.650	0.878	<u>0.805</u>	<u>0.636</u>

Table 4. Comparison of *SG-ESSIM* to several other state-of-the-art algorithms on TID2008 [21] and CSIQ [8]. The highest values are typed in **bold**, while the second highest ones are underlined.

FR-IQA Metric	TID2008 [21]			CSIQ [8]		
	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC
2stepQA [50]	0.757	0.769	0.574	0.841	0.849	0.655
CSV [51]	<u>0.852</u>	0.851	0.659	<u>0.933</u>	0.933	0.766
DISTS [52]	0.705	0.668	0.488	0.930	0.930	0.764
ESSIM [19]	0.658	<u>0.876</u>	<u>0.696</u>	0.814	0.933	0.768
GSM [29]	0.782	0.781	0.578	0.906	0.910	0.729
IW-SSIM [16]	0.842	0.856	0.664	0.804	0.921	0.753
MAD [8]	0.831	0.829	0.639	0.950	0.947	0.796
MS-SSIM [26]	0.838	0.846	0.648	0.913	0.917	0.743
PSNR	0.447	0.489	0.346	0.853	0.809	0.599
ReSIFT [53]	0.627	0.632	0.484	0.884	0.868	0.695
RVSIM [54]	0.789	0.743	0.566	0.923	0.903	0.728
SSIM [23]	0.669	0.675	0.485	0.812	0.812	0.606
SUMMER [55]	0.817	0.823	0.623	0.826	0.830	0.658
<i>SG-ESSIM</i>	0.853	0.888	0.708	0.836	<u>0.944</u>	<u>0.786</u>

Table 5. Comparison of *SG-ESSIM* to several other state-of-the-art algorithms. Direct and weighted average PLCC, SROCC, and KROCC values are presented. Measured on KADID-10k [20], TID2013 [7], TID2008 [21], and CSIQ [8]. The highest values are typed in **bold**, while the second highest ones are underlined.

FR-IQA Metric	Direct Average			Weighted Average		
	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC
2stepQA [50]	0.776	0.781	0.587	0.765	0.768	0.572
CSV [51]	0.827	0.825	0.653	0.740	0.738	0.582
DISTS [52]	0.801	0.781	0.601	0.795	0.785	0.599
ESSIM [19]	0.714	<u>0.857</u>	<u>0.681</u>	0.673	<u>0.830</u>	<u>0.647</u>
GSM [29]	0.814	0.815	0.622	0.789	0.789	0.596
IW-SSIM [16]	0.815	0.828	0.635	<u>0.800</u>	0.780	0.570
MAD [8]	<u>0.831</u>	0.820	0.643	0.763	0.758	0.573
MS-SSIM [26]	0.841	0.842	0.656	0.821	0.822	0.633
PSNR	0.599	0.655	0.475	0.520	0.660	0.470
ReSIFT [53]	0.697	0.688	0.530	0.655	0.641	0.483
RVSIM [54]	0.801	0.762	0.589	0.752	0.725	0.549
SSIM [23]	0.692	0.694	0.504	0.668	0.669	0.485
SUMMER [55]	0.746	0.750	0.573	0.720	0.720	0.540
<i>SG-ESSIM</i>	0.827	0.869	0.695	0.783	0.843	0.661

Table 6. Comparison of TID2013's [7] distortion types. SROCC values are given. The highest values are typed in **bold**, while the second highest ones are underlined.

	2stepQA [50]	CSV [51]	DISTS [52]	ESSIM [19]	GSM [29]	MAD [8]	MS-SSIM [26]	ReSIFT [53]	RVSIM [54]	SSIM [23]	SG-ESSIM
AGN	0.817	0.938	0.845	0.911	0.899	0.912	0.624	0.831	0.886	0.848	<u>0.936</u>
ANC	0.590	0.862	0.786	0.806	0.823	0.800	0.387	0.749	0.836	0.779	<u>0.855</u>
SCN	0.860	0.939	0.859	<u>0.938</u>	0.927	0.929	0.683	0.839	0.868	0.851	0.935
MN	0.395	0.748	0.814	0.711	0.704	0.658	0.372	0.702	0.734	<u>0.775</u>	0.715
HFN	0.828	0.927	0.868	0.890	0.884	0.902	0.704	0.869	0.895	0.889	<u>0.920</u>
IN	0.715	<u>0.848</u>	0.674	0.825	0.813	0.743	0.766	0.824	0.865	0.810	0.833
QN	0.886	0.892	0.810	0.904	0.911	<u>0.895</u>	0.720	0.745	0.869	0.817	0.911
GB	0.853	0.933	0.926	0.970	0.954	0.915	0.762	0.937	0.970	0.910	<u>0.969</u>
DEN	0.900	0.952	0.899	<u>0.956</u>	0.955	0.922	0.819	0.907	0.926	0.876	0.963
JPEG	0.867	<u>0.944</u>	0.897	0.923	0.933	0.924	0.784	0.905	0.930	0.893	0.950
JP2K	0.891	0.966	0.931	0.946	0.934	0.929	0.790	0.928	0.946	0.806	<u>0.949</u>
JGTE	0.806	0.800	0.906	0.826	<u>0.866</u>	0.768	0.582	0.712	0.831	0.701	0.823
J2TE	0.854	0.887	0.865	0.902	0.893	0.854	0.742	0.835	0.882	0.813	<u>0.899</u>
NEPN	0.775	<u>0.811</u>	0.833	0.799	0.804	0.803	0.792	0.693	0.771	0.634	0.801
BLOCK	0.044	0.183	0.302	0.649	0.588	−0.322	0.382	0.440	0.545	0.564	<u>0.623</u>
MS	0.660	0.654	0.752	0.712	0.728	0.708	0.732	0.418	0.559	<u>0.738</u>	0.706
CC	0.430	0.227	<u>0.464</u>	0.453	0.466	0.420	0.027	−0.055	0.132	0.355	0.452
CCS	−0.258	0.809	<u>0.789</u>	−0.297	0.676	−0.059	−0.055	−0.209	0.366	0.742	0.010
MGN	0.747	0.884	0.790	0.853	0.831	<u>0.888</u>	0.653	0.765	0.853	0.804	0.900
CN	0.858	0.924	0.907	0.910	0.902	<u>0.904</u>	0.596	0.882	0.914	0.797	<u>0.916</u>
LCNI	0.902	0.965	0.932	0.957	0.945	0.950	0.713	0.897	0.933	0.877	<u>0.952</u>
ICQD	0.808	<u>0.919</u>	0.832	0.904	0.901	0.867	0.739	0.770	0.871	0.820	0.928
CA	0.702	0.845	0.879	0.839	0.835	0.760	0.568	0.838	<u>0.871</u>	0.740	0.835
SSR	0.926	0.976	0.944	<u>0.965</u>	0.961	0.949	0.801	0.944	<u>0.956</u>	0.822	0.964
All	0.733	0.848	0.711	0.797	0.787	0.778	0.785	0.623	0.683	0.616	<u>0.805</u>

Table 7. Comparison of TID2008’s [21] distortion types. SROCC values are given. The highest values are typed in **bold**, while the second highest ones are underlined.

	2stepQA [50]	CSV [51]	DISTS [52]	ESSIM [19]	GSM [29]	MAD [8]	MS-SSIM [26]	ReSIFT [53]	RVSIM [54]	SSIM [23]	SG-ESSIM
AGN	0.766	0.922	0.812	0.875	0.855	0.872	0.610	0.771	0.840	0.805	0.913
ANC	0.627	<u>0.893</u>	0.811	0.792	0.821	0.803	0.354	0.762	0.829	0.780	0.900
SCN	0.814	0.932	0.838	0.909	0.904	0.901	0.727	0.810	0.837	0.800	0.920
MN	0.450	0.781	0.830	0.744	0.736	0.673	0.304	0.728	0.760	0.797	<u>0.825</u>
HFN	0.818	0.936	0.870	0.899	0.889	0.894	0.749	0.881	0.886	0.871	0.921
IN	0.659	<u>0.819</u>	0.626	0.777	0.764	0.650	0.767	0.777	0.836	0.776	0.786
QN	0.850	<u>0.894</u>	0.770	0.884	0.903	0.851	0.708	0.730	0.836	0.784	0.873
GB	0.877	<u>0.923</u>	0.909	0.966	0.948	0.896	0.759	0.904	0.963	0.866	<u>0.964</u>
DEN	0.919	<u>0.970</u>	0.931	0.974	<u>0.971</u>	0.928	0.786	0.923	0.939	0.873	0.963
JPEG	0.895	<u>0.948</u>	0.894	0.938	<u>0.937</u>	0.931	0.774	0.914	0.926	0.880	0.959
JP2K	0.910	0.984	0.953	0.966	0.949	0.941	0.837	0.935	0.970	0.745	<u>0.972</u>
JGTE	0.851	0.790	0.907	0.859	<u>0.871</u>	0.781	0.606	0.735	0.860	0.666	<u>0.855</u>
J2TE	0.845	0.852	0.833	0.875	0.880	0.802	0.742	0.778	0.854	0.769	0.863
NEPN	<u>0.803</u>	0.752	0.882	0.742	0.784	0.801	0.749	0.761	0.732	0.588	0.729
Block	0.441	0.770	0.618	<u>0.876</u>	0.843	−0.362	0.765	0.743	0.782	0.804	0.905
MS	0.655	0.594	<u>0.681</u>	0.611	0.638	0.563	0.711	0.322	0.525	0.629	0.683
CC	0.597	0.330	0.649	0.624	0.634	0.548	0.042	−0.018	0.194	0.502	<u>0.642</u>
All	0.769	0.851	0.668	<u>0.876</u>	0.781	0.829	0.846	0.632	0.743	0.675	0.888

Table 8. Comparison of SROCC of each FR-IQA metrics on TID2013’s [7] distortion levels (Level 1 represents the lowest level of degradation, while Level 5 represents the highest one). The highest values are typed in **bold**, while the second highest ones are underlined.

	2stepQA [50]	CSV [51]	DISTS [52]	ESSIM [19]	GSM [29]	MAD [8]	MS-SSIM [26]	ReSIFT [53]	RVSIM [54]	SSIM [23]	SG-ESSIM
Level 1	0.246	<u>0.424</u>	0.235	0.388	0.372	0.388	0.166	0.181	0.248	0.204	0.448
Level 2	0.394	0.626	0.440	0.547	0.512	0.368	0.049	0.401	0.430	0.276	<u>0.569</u>
Level 3	0.539	0.635	0.367	<u>0.638</u>	0.523	0.442	0.240	0.415	0.416	0.084	0.660
Level 4	0.571	0.749	0.606	<u>0.766</u>	0.669	0.284	0.172	0.699	0.702	0.208	0.787
Level 5	0.663	0.787	0.664	0.875	0.745	0.308	0.397	0.788	0.803	0.202	<u>0.861</u>
All	0.733	0.848	0.711	0.797	0.787	0.778	0.785	0.623	0.683	0.616	<u>0.805</u>

Table 9. Comparison of SROCC for each FR-IQA metric on TID2008’s [21] distortion levels (Level 1 represents the lowest level of degradation, while Level 5 represents the highest one). The highest values are typed in **bold**, while the second highest ones are underlined.

	2stepQA [50]	CSV [51]	DISTS [52]	ESSIM [19]	GSM [29]	MAD [8]	MS-SSIM [26]	ReSIFT [53]	RVSIM [54]	SSIM [23]	SG-ESSIM
Level 1	0.470	0.638	0.566	0.655	0.639	0.432	0.067	0.457	0.634	0.368	0.691
Level 2	0.619	0.683	0.381	<u>0.773</u>	0.636	0.520	0.221	0.437	0.513	0.105	0.807
Level 3	0.573	0.774	0.581	<u>0.826</u>	0.677	0.239	0.059	0.707	0.761	0.190	0.849
Level 4	0.610	0.829	0.628	0.905	0.718	0.232	0.275	0.788	0.825	0.241	<u>0.891</u>
All	0.769	0.851	0.668	<u>0.876</u>	0.781	0.829	0.846	0.632	0.743	0.675	0.888

In Figure 3, the execution time (logarithmic scale) versus SROCC scatter plot measured on KADID-10k [20] is depicted. From this figure, it could be seen that the proposed *SG-ESSIM* method was the fourth fastest algorithm out of the examined thirteen ones. Considering execution time and estimation performance together, the *SG-ESSIM* provided a competitive result against the state-of-the-art. Similarly, Figure 4 depicts the execution versus the SROCC on the CSIQ [8] database. Here, we could also observe that considering execution time and performance together, the proposed method was able to achieve a competitive result.

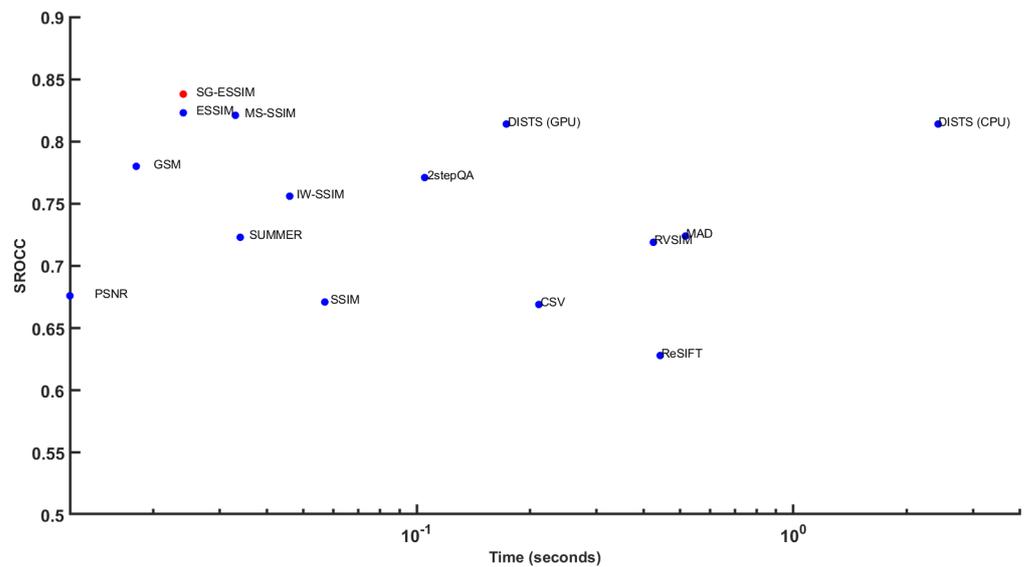


Figure 3. Execution time (logarithmic scale) versus SROCC measured on KADID-10k [20].

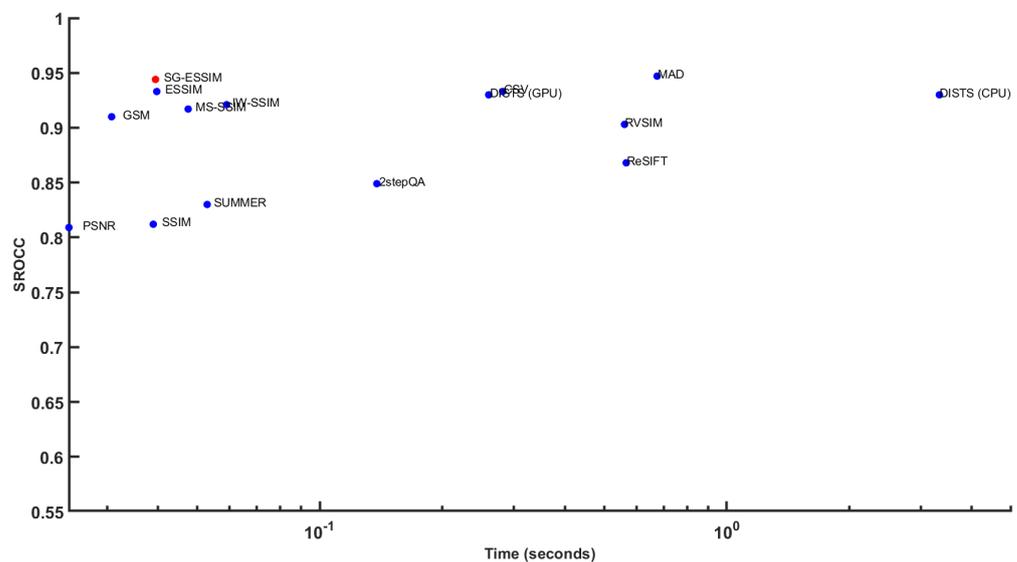


Figure 4. Execution time (logarithmic scale) versus SROCC measured on CSIQ [8].

6. Conclusions

The goal of FR-IQA is to predict digital images’ perceptual quality with full access to the distortion-free, reference images. FR-IQA methods usually contain two stages: local image quality estimation and pooling. In addition, visual saliency is utilized in the pooling stage as weights for the weighted averaging of local image quality. Unlike this traditional approach, we applied visual saliency in the computation of local image quality, motivated by the fact that local image quality is determined both by local image degradation and

visual saliency simultaneously. Experimental results on four publicly available benchmark IQA databases showed that the proposed algorithm was able to outperform the state-of-the-art and was characterized by low computational costs. Future work could involve the incorporation of local and global saliency into a novel FR-IQA metric. The source code of the proposed method is available at: <https://github.com/Skythianos/SG-ESSIM> (accessed on 12 June 2022).

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: In this paper, the following publicly available benchmark databases were used: 1. KADID-10k: <http://database.mmsp-kn.de/kadid-10k-database.html> (accessed on 12 May 2022), 2. TID2013: <http://www.ponomarenko.info/tid2013.htm> (accessed on 12 May 2022), 3. TID2008: <http://www.ponomarenko.info/tid2008.htm> (accessed on 12 May 2022), 4. CSIQ: https://isp.uv.es/data_quality.html (accessed on 12 May 2022).

Acknowledgments: We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations were used in this manuscript:

AGN	additive Gaussian noise;
ANC	additive noise in color components;
CA	chromatic aberrations;
CC	contrast change;
CCS	change in color saturation;
CN	comfort noise;
CNN	convolutional neural network;
CPU	central processing unit;
DEN	image denoising;
FR-IQA	full-reference image quality assessment;
GB	Gaussian blur;
GPU	graphics processing unit;
HFN	high frequency noise;
HVS	human visual system;
ICQD	image color quantization with dither;
IN	impulse noise;
IQA	image quality assessment;
JGTE	JPEG transmission error;
JPEG	joint photographic experts group;
KROCC	Kendall's rank order correlation coefficient;
LCNI	lossy compression of noisy images;
MGN	multiplicative Gaussian noise;
MN	masked noise;
MOS	mean opinion score;
MS	mean shift;
MSE	mean squared error;
NEPN	noneccentricity pattern noise;
NR-IQA	no-reference image quality assessment;
PLCC	Pearson's linear correlation coefficient;
PSNR	peak signal-to-noise ratio;
QN	quantization noise;
RR-IQA	reduced-reference image quality assessment;

SCN	spatially correlated noise;
SG	saliency guided;
SROCC	Spearman's rank order correlation coefficient;
SSR	space sampling and reconstruction

References

- Ding, K.; Ma, K.; Wang, S.; Simoncelli, E.P. Comparison of full-reference image quality models for optimization of image processing systems. *Int. J. Comput. Vis.* **2021**, *129*, 1258–1281. [[CrossRef](#)]
- Chen, B.; Zhu, L.; Zhu, H.; Yang, W.; Lu, F.; Wang, S. The Loop Game: Quality Assessment and Optimization for Low-Light Image Enhancement. *arXiv* **2022**, arXiv:2202.09738.
- Goyal, B.; Gupta, A.; Dogra, A.; Koundal, D. An adaptive bitonic filtering based edge fusion algorithm for Gaussian denoising. *Int. J. Cogn. Comput. Eng.* **2022**, *3*, 90–97. [[CrossRef](#)]
- Saito, Y.; Miyata, T. Recovering Texture with a Denoising-Process-Aware LMMSE Filter. *Signals* **2021**, *2*, 286–303. [[CrossRef](#)]
- Chubarau, A.; Akhavan, T.; Yoo, H.; Mantiuk, R.K.; Clark, J. Perceptual image quality assessment for various viewing conditions and display systems. *Electron. Imaging* **2020**, *2020*, 67-1. [[CrossRef](#)]
- Saupe, D.; Hahn, F.; Hosu, V.; Zingman, I.; Rana, M.; Li, S. Crowd workers proven useful: A comparative study of subjective video quality assessment. In Proceedings of the QoMEX 2016: 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 6–8 June 2016.
- Ponomarenko, N.; Jeremeiev, O.; Lukin, V.; Egiazarian, K.; Jin, L.; Astola, J.; Vozel, B.; Chehdi, K.; Carli, M.; Battisti, F.; et al. Color image database TID2013: Peculiarities and preliminary results. In Proceedings of the European Workshop on Visual Information Processing (EUVIP), Paris, France, 10–12 June 2013; pp. 106–111.
- Larson, E.C.; Chandler, D.M. Most apparent distortion: Full-reference image quality assessment and the role of strategy. *J. Electron. Imaging* **2010**, *19*, 011006.
- Zhai, G.; Min, X. Perceptual image quality assessment: A survey. *Sci. China Inf. Sci.* **2020**, *63*, 1–52. [[CrossRef](#)]
- Liu, H.; Heynderickx, I. Visual attention in objective image quality assessment: Based on eye-tracking data. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *21*, 971–982.
- Zhang, L.; Shen, Y.; Li, H. VSI: A visual saliency-induced index for perceptual image quality assessment. *IEEE Trans. Image Process.* **2014**, *23*, 4270–4281. [[CrossRef](#)]
- Zhang, W.; Borji, A.; Wang, Z.; Le Callet, P.; Liu, H. The application of visual saliency models in objective image quality assessment: A statistical evaluation. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 1266–1278. [[CrossRef](#)]
- Ma, Q.; Zhang, L. Saliency-based image quality assessment criterion. In Proceedings of the International Conference on Intelligent Computing, Shanghai, China, 15–18 September 2008; pp. 1124–1133.
- Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)]
- Kovesi, P.; Robotics and Vision Research Group. Image features from phase congruency. *Videre J. Comput. Vis. Res.* **1999**, *1*, 1–26.
- Wang, Z.; Li, Q. Information content weighting for perceptual image quality assessment. *IEEE Trans. Image Process.* **2010**, *20*, 1185–1198. [[CrossRef](#)]
- Shi, C.; Lin, Y. Full reference image quality assessment based on visual salience with color appearance and gradient similarity. *IEEE Access* **2020**, *8*, 97310–97320. [[CrossRef](#)]
- Varga, D. Full-Reference Image Quality Assessment Based on Grünwald–Letnikov Derivative, Image Gradients, and Visual Saliency. *Electronics* **2022**, *11*, 559. [[CrossRef](#)]
- Zhang, X.; Feng, X.; Wang, W.; Xue, W. Edge strength similarity for image quality assessment. *IEEE Signal Process. Lett.* **2013**, *20*, 319–322. [[CrossRef](#)]
- Lin, H.; Hosu, V.; Saupe, D. KADID-10k: A large-scale artificially distorted IQA database. In Proceedings of the 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 5–7 June 2019; pp. 1–3.
- Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F. TID2008-a database for evaluation of full-reference visual quality assessment metrics. *Adv. Mod. Radioelectron.* **2009**, *10*, 30–45.
- Yang, X.; Sun, Q.; Wang, T. Image quality assessment via spatial structural analysis. *Comput. Electr. Eng.* **2018**, *70*, 349–365. [[CrossRef](#)]
- Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
- Brunet, D.; Vrscay, E.R.; Wang, Z. On the mathematical properties of the structural similarity index. *IEEE Trans. Image Process.* **2011**, *21*, 1488–1499. [[CrossRef](#)]
- Nilsson, J.; Akenine-Möller, T. Understanding ssim. *arXiv* **2020**, arXiv:2006.13846.
- Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, Pacific Grove, CA, USA, 9–12 November 2003; Volume 2, pp. 1398–1402.
- Sampat, M.P.; Wang, Z.; Gupta, S.; Bovik, A.C.; Markey, M.K. Complex wavelet structural similarity: A new image similarity index. *IEEE Trans. Image Process.* **2009**, *18*, 2385–2401. [[CrossRef](#)]

28. Chen, G.H.; Yang, C.L.; Po, L.M.; Xie, S.L. Edge-based structural similarity for image quality assessment. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, Toulouse, France, 14–19 May 2006; Volume 2, p. II.
29. Liu, A.; Lin, W.; Narwaria, M. Image quality assessment based on gradient similarity. *IEEE Trans. Image Process.* **2011**, *21*, 1500–1512. [[PubMed](#)]
30. Zhu, J.; Wang, N. Image quality assessment by visual gradient similarity. *IEEE Trans. Image Process.* **2011**, *21*, 919–933. [[PubMed](#)]
31. Ma, J.; Wu, J.; Li, L.; Dong, W.; Xie, X.; Shi, G.; Lin, W. Blind image quality assessment with active inference. *IEEE Trans. Image Process.* **2021**, *30*, 3650–3663. [[CrossRef](#)]
32. Chetouani, A.; Pedersen, M. Image Quality Assessment without Reference by Combining Deep Learning-Based Features and Viewing Distance. *Appl. Sci.* **2021**, *11*, 4661. [[CrossRef](#)]
33. Vosta, S.; Yow, K.C. A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras. *Appl. Sci.* **2022**, *12*, 1021. [[CrossRef](#)]
34. Amirshahi, S.A.; Pedersen, M.; Yu, S.X. Image quality assessment by comparing CNN features between images. *J. Imaging Sci. Technol.* **2016**, *60*, 60410-1. [[CrossRef](#)]
35. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
36. Barla, A.; Franceschi, E.; Odone, F.; Verri, A. Image kernels. In Proceedings of the International Workshop on Support Vector Machines, Niagara Falls, ON, Canada, 10 August 2002; pp. 83–96.
37. Amirshahi, S.A.; Pedersen, M.; Beghdadi, A. Reviving traditional image quality metrics using CNNs. *Color Imaging Conf.* **2018**, *26*, 241–246. [[CrossRef](#)]
38. Ahn, S.; Choi, Y.; Yoon, K. Deep learning-based distortion sensitivity prediction for full-reference image quality assessment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 344–353.
39. Chubarau, A.; Clark, J. VTAMIQ: Transformers for Attention Modulated Image Quality Assessment. *arXiv* **2021**, arXiv:2110.01655.
40. Bosse, S.; Maniry, D.; Müller, K.R.; Wiegand, T.; Samek, W. Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Trans. Image Process.* **2017**, *27*, 206–219. [[CrossRef](#)] [[PubMed](#)]
41. Okarma, K. Combined full-reference image quality metric linearly correlated with subjective assessment. In Proceedings of the International Conference on Artificial Intelligence and Soft Computing, Zakopane, Poland, 13–17 June 2010; pp. 539–546.
42. Okarma, K. Combined image similarity index. *Opt. Rev.* **2012**, *19*, 349–354. [[CrossRef](#)]
43. Okarma, K. Extended hybrid image similarity–combined full-reference image quality metric linearly correlated with subjective scores. *Elektron. Elektrotehnika* **2013**, *19*, 129–132. [[CrossRef](#)]
44. Oszust, M. Full-reference image quality assessment with linear combination of genetically selected quality measures. *PLoS ONE* **2016**, *11*, e0158333. [[CrossRef](#)] [[PubMed](#)]
45. Lukin, V.V.; Ponomarenko, N.N.; Ieremeiev, O.I.; Egiazarian, K.O.; Astola, J. Combining full-reference image visual quality metrics by neural network. In Proceedings of the Human Vision and Electronic Imaging XX, SPIE, San Francisco, CA, USA, 9–12 February 2015; Volume 9394, pp. 172–183.
46. Levkine, G. Prewitt, Sobel and Scharr gradient 5×5 convolution matrices. *Image Process. Artic. Second. Draft.* **2012**, 1–17.
47. Kim, D.J.; Lee, H.C.; Lee, T.S.; Lee, K.W.; Lee, B.H. Edge-Based Gaze Planning for Salient Proto-Objects. *Appl. Mech. Mater.* **2013**, *330*, 1003–1007. [[CrossRef](#)]
48. Pedersen, M.; Hardeberg, J.Y. Full-reference image quality metrics: Classification and evaluation. *Found. Trends[®] Comput. Graph. Vis.* **2012**, *7*, 1–80.
49. Sheikh, H.R.; Bovik, A.C.; De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **2005**, *14*, 2117–2128. [[CrossRef](#)]
50. Yu, X.; Bampis, C.G.; Gupta, P.; Bovik, A.C. Predicting the quality of images compressed after distortion in two steps. *IEEE Trans. Image Process.* **2019**, *28*, 5757–5770. [[CrossRef](#)]
51. Temel, D.; AlRegib, G. CSV: Image quality assessment based on color, structure, and visual system. *Signal Process. Image Commun.* **2016**, *48*, 92–103. [[CrossRef](#)]
52. Ding, K.; Ma, K.; Wang, S.; Simoncelli, E.P. Image quality assessment: Unifying structure and texture similarity. *arXiv* **2020**, arXiv:2004.07728.
53. Temel, D.; AlRegib, G. ReSIFT: Reliability-weighted sift-based image quality assessment. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 2047–2051.
54. Yang, G.; Li, D.; Lu, F.; Liao, Y.; Yang, W. RVSIM: A feature similarity method for full-reference image quality assessment. *EURASIP J. Image Video Process.* **2018**, *2018*, 1–15. [[CrossRef](#)]
55. Temel, D.; AlRegib, G. Perceptual image quality assessment through spectral analysis of error representations. *Signal Process. Image Commun.* **2019**, *70*, 37–46. [[CrossRef](#)]