

Review

Not peer-reviewed version

Overview of High Dynamic Range Image Quality Assessment

[Yue Liu](#) , [Yu Tian](#) , Shiqi Wang , Xinfeng Zhang , [Sam Tak Wu Kwong](#) *

Posted Date: 13 August 2024

doi: 10.20944/preprints202408.0935.v1

Keywords: High dynamic range; Objective Image quality assessment; Subjective Image quality 10 assessment



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Overview of High Dynamic Range Image Quality Assessment

Yue Liu ^{1,†} , Yu Tian ^{1,†} , Shiqi Wang ¹ , Xinfeng Zhang ²  and Sam Kwong ^{3,*} 

¹ Department of Computer Science, City University of Hong Kong; Hong Kong; yliu724-c@my.cityu.edu.hk (Y.L.); ytian73-c@my.cityu.edu.hk (Y.T.); shiqiwan@cityu.edu.hk (S.W.)

² School of Computer Science and Technology, University of Chinese Academic of Sciences, Beijing, China; xfzhang@ucas.ac.cn (X.Z.)

³ Lingnan University, Hong Kong; samkwong@ln.edu.hk (S.K.)

* Correspondence: samkwong@ln.edu.hk

† These authors contributed equally to this work.

Abstract: In recent years, the High Dynamic Range (HDR) image has gained widespread popularity across various domains such as security, multimedia, and biomedical fields, owing to its ability to deliver an authentic visual experience. However, the extensive dynamic range and rich detail in HDR images present challenges in assessing their quality. Therefore, current efforts involve constructing subjective databases and proposing objective quality assessment metrics to achieve efficient HDR Image Quality Assessment (IQA). Recognizing the absence of a systematic overview of these approaches, this paper provides a comprehensive survey of existing HDR IQA methods. Furthermore, this paper aims to serve as a valuable resource for researchers by discussing the limitations of current methodologies and potential research directions in the future.

Keywords: high dynamic range; objective image quality assessment; subjective image quality assessment

1. Introduction

Compared to ordinary Low Dynamic Range (LDR) images, High Dynamic Range (HDR) images exhibit a wider dynamic range (i.e., a larger ratio between the highest and lowest value in the image), resulting in superior overall contrast effects and richer detail, especially in the extremely bright or dark regions. Due to its capability to offer highly realistic visual experiences to viewers, recent years have witnessed a significant increase in the research of HDR images, including photography [1], automotive [2], medical imaging [3], remote sensing [4], etc. In this context, the HDR Image Quality Assessment (IQA) becomes a fundamental issue in the field of HDR image processing tasks, which can be widely employed for optimizing parameters in various codecs and learning-based models. However, it is not suitable to directly apply LDR IQA metrics to HDR images due to the substantial disparities in statistical distribution and luminance range between LDR and HDR images. Therefore, specialized evaluation methods and metrics tailored for HDR images are essential to assess their quality effectively.

In general, existing IQA algorithms can be classified into subjective and objective methods according to the evaluation approach. With the explosive growth of HDR content, the pressing issue at hand pertains to the storage and transmission of HDR images. Consequently, existing subjective experiments [5–7] are dedicated to exploring the impact of compression algorithms on the visual quality of HDR images. Nevertheless, the subjective experiment is usually time-consuming and expensive. In contrast, objective methods are considered more efficient due to their ability to directly extract features that closely align with the Human Visual System (HVS) from the images. Herein, the HDR-VDP family [8–10] propose to model the luminance perceived by HVS. Some methods propose to evaluate the quality of HDR images in the perceptual domain [11,12], while others adapt existing metrics to predict quality scores [13,14]. Despite great progress, HDR IQA still encounters some challenges, particularly in the lack of large subjective databases, as well as how to properly compress the dynamic range of the HDR image into the perceptual domain.

The necessity for a comprehensive survey of HDR IQA methods arises from the limitations of existing surveys, which fail to cover the scope of the field comprehensively or lack summaries of the latest methodologies [15]. Therefore, this paper aims to fill this gap by providing an extensive overview of existing HDR IQA algorithms, including both subjective and objective methods. Furthermore, we discuss the limitations of current subjective and objective HDR IQA approaches and potential directions for improvement, hoping to provide comprehensive insight for future research.

The structure of this paper is outlined as follows: Section 2 introduces existing subjective HDR IQA databases, including detailed descriptions of the database and experimental settings. Section 3 summarizes existing objective HDR IQA algorithms. The limitations of both subjective and objective HDR IQA approaches are discussed in Section 4, together with the potential directions for future research.

2. Subjective HDR Image Quality Assessment Databases

Existing subjective HDR IQA databases focus on the evaluation of the visual quality of HDR images with compression distortions. In contrast to LDR images that are typically encoded in 8-bit integers, the most commonly used formats for HDR images are Radiance RGBE and OpenEXR, where the data are stored in 16-bit or 32-bit floating-point numbers. Therefore, traditional image compression schemes like JPEG and JPEG2000 designed for LDR images can not handle the compression of HDR images directly. Due to the widespread application of LDR codecs, converting the HDR image to the LDR image to make it compatible with existing codecs is a highly intuitive approach with practical applications. As shown in Figure 1, the existing compression scheme [16,17] for HDR images can be summarized into three steps. In the pre-processing stage, the input HDR image is converted to an LDR image by a Tone Mapping Operator (TMO). Herein, Sugiyama *et al.* propose to improve the quality of the tone-mapped LDR image by an optimization module [17]. Afterward, the LDR image is fed into the codec (JPEG or JPEG2000), together with the additional information such as the parameters for inverse TMO (iTMO), or the different information between the HDR and corresponding LDR image. In the post-preprocessing stage, the content to be displayed is alternative according to the conditions of the terminal monitor. If the monitor supports HDR content, the HDR image will be generated from the decoded LDR image by the iTMO with additional information. Otherwise, the decoded LDR image can be directly displayed. With such architecture, the visual quality of the compressed HDR images has been evaluated in existing subjective databases, which are summarized in Table 1 and described as follows.

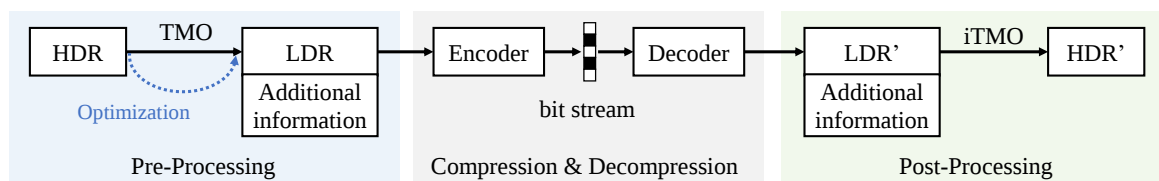


Figure 1. Overview of the HDR image compression schemes.

To evaluate the performance of 1) optimization methods in HDR compression and 2) the objective IQA algorithms for compressed HDR images, Narwaria *et al.* generate a database [5] and conduct a subjective experiment to obtain the corresponding mean opinion score (MOS). The database is composed of 10 reference HDR images and 140 distorted HDR images, which are compressed with the JPEG scheme with 7 bitrates and 2 optimization methods. The Mean Square Error (MSE) and Structural Similarity Index (SSIM) [18] are used to optimize the quality of the generated LDR image with the iCAM06 [19] TMO. The quality scores of the 150 HDR images are collected from 27 observers using the absolute category rating with hidden reference (ACR-HR) methodology [20]. The subjective experiment is conducted in a room with 130 cd/m^2 luminance condition. The HDR images are displayed on a SIM2 Solar47 HDR monitor [21] with a maximum luminance value of up to 4000 cd/m^2 .

From the experimental results, the SSIM does not outperform MSE when they are used to optimize the quality of tone-mapped LDR images. Furthermore, the HDR-VDP-2 [9] shows better performance on the overall results. However, when it comes to statistical analysis, the HDR-VDP-2 is indistinguishable from the SIQM method [22].

Table 1. Overview of existing HDR IQA databases. The observers, reference images, distorted images, distortion type, and subjective methodology are denoted as Obs., Ref., Dis., Dist., and Meth., respectively.

database	#Obs.	#Ref.	#Dis.	Dist.	TMO	Meth.
Narwaria2013 [5]	27	10	140	JPEG	iCAM06	ACR-HR
Narwaria2014 [23]	29	6	210	JPEG2000	AL, Dur RG, RL Log	ACR-HR
Korshunov2015 [6]	24	20	240	JPEG-Xt	RG MT [24]	DSIS
Valenzise2014 [7]	15	5	50	JPEG JPEG2000 JPEG-Xt	Mai	DSIS
Zerman2017 [25]	15	5	50	JPEG JPEG2000	Mai, PQ	DSIS
UPIQ [26] (HDR part)	20	30	380	JPEG JPEG-Xt	iCAM06 RG MT [24]	-
HDRC [27]	20	80	400	JPEG-Xt VVC	RG PQ	DSIS

In the database proposed by Narwaria *et al.* [23], the performance of five TMOs with the JPEG2000 compression scheme is subjectively evaluated. Herein, 3 local TMOs proposed by Reinhard (RL) [28], Durand (Dur) [29], and Ashikmin (AL) [30], together with 2 global TMOs, logarithmic TMO (Log) and the global TMO (RG) proposed in [28], are applied in the experiments. In this database, 6 reference HDR scenes are selected to obtain 210 distorted HDR images with 7-bit rates for each TMO. The subjective experiment settings are the same as the conditions in [5]. The MOS values of the 216 HDR images are obtained from 29 participants with ACR-HR methodology. Experimental analysis results reveal that the AL TMO statistically performs better than other TMOs. Furthermore, the local TMOs can help to achieve better visual quality of the compressed HDR images, while the performance of the two global TMOs does not show statistical differences. Korsunov *et al.* [6] compress the HDR image with the JPEG-Xt [31] compression scheme and conduct the subjective experiment to build an HDR IQA database. In this database, 20 reference HDR images with various content are cropped and adjusted to the size of 944×1080 for display. 240 compressed HDR images are generated by applying 3 profiles and 4-bit rates for each scene. The subjective experiment is conducted in a room with 20 cd/m^2 background luminance value. A SIM2 HDR monitor with luminance values varying from 0.001 to 4000 cd/m^2 is adopted to display the HDR images. The double-stimulus impairment scale (DSIS) methodology [32] is applied to acquire the MOS values from 24 subjects. Valenzise *et al.* [7] have developed a database specifically for evaluating existing full-reference HDR IQA metrics. The database is composed of 5 reference HDR images and 50 distorted HDR images, where the JPEG, JPEG2000, and JPEG XT schemes are used for image compression. The TMO used in this database is proposed by Mai *et al.* [33]. The subjective experiment is conducted in a room with 20 cd/m^2 background luminance value. The HDR images are presented on a SIM2 HDR47 monitor. The MOS scores are collected by the DSIS methodology from 15 subjects, where the continuous quality scores from 0 to 100 are divided into five scales, denoting the distortion level from “very annoying” to “Imperceptible”. Based on this

database, they evaluate the performance of the objective IQA methods, including HDR-VDP-2 [9] tailored for HDR images, alongside PSNR and SSIM, which were adapted for the assessment of perceptual mapped HDR images by logarithmic function or perceptual uniform (PU) encoding [34]. From the results, the PU-SSIM outperforms other metrics, followed by the log-SSIM and HDR-VDP-2 algorithms.

In order to conduct a more comprehensive assessment of the performance of full-reference HDR IQA metrics, Zerman *et al.* [25] introduce a novel database building upon their previous work [7]. This database comprises 5 reference HDR images and 50 HDR images distorted by the JPEG and JPEG2000 compression schemes. Herein, two TMOs are employed to generate the LDR images: the TMO proposed by Mai *et al.* [24], and the Perceptual Quantizer (PQ) [35] published in SMPTE ST 2084. The subjective experiment settings are controlled the same as in Valenzise2014 [7]. Subsequently, the obtained MOS values are further aligned with the MOS values in [7] for further analysis. In the experiment, 25 IQA metrics are evaluated on the proposed database and 4 existing databases. There are three interesting conclusions can be drawn from the results. Firstly, the Narwaria2014 [23] is a much more challenging database due to its complex distortion types. Secondly, metrics tailored for HDR content, such as HDR-VDP-2.2 [36] and HDR-VQM [11], exhibit superior performance in comparison to other IQA metrics. Nevertheless, the probability of misclassification still remains as high as 20%, indicating the imperative for the design of HDR IQA metrics that better align with HVS. Furthermore, concerning compression distortions, structural loss in the luminance domain is more perceptible to the HVS than loss in the color space. Due to the limited number of images in the previous HDR IQA databases, Mikhailiuk *et al.* [26] construct a Unified Photometric Image Quality (UPIQ) database with 380 HDR images and 3779 LDR images. The HDR part is collected from the two databases Korsunov2015 [6] and Narwaria2013 [5]. The MOS values are aligned by the psychometric scaling [37] methodology. Recently, Liu *et al.* [27] proposed an HDR compression (HRDC) database, with the goal of examining the perceptual quality of HDR images compressed using the JPEG-Xt and VVC codecs. The experiment is conducted using a Samsung QN90A TV, with 20 participants providing opinion scores based on the DSIS testing methodology. This database is the largest open-source database available for the HDR IQA task, including 400 distorted HDR images with JPEG-Xt and VVC compression and 80 reference HDR images. Based on this database, they conduct extensive experiments to evaluate the effectiveness of both HDR and LDR IQA metrics on the HDR IQA task. Furthermore, they also evaluate the effects of different perceptual mapping methods, such as PQ, PU, and PU21 mapping methods, on the performance of LDR IQA metrics. Based on the experimental results, these perceptual mapping methods exhibit significant variations in performance across different databases. Moreover, no single perceptual mapping method achieves optimal performance across all databases, underscoring the challenges posed by HDR image evaluation and suggesting the need for new mapping methods to achieve stable and effective quality assessment for HDR images.

3. Objective HDR Image Quality Assessment Methods

Recent years have witnessed significant development of efficient objective LDR IQA metrics [38–45], which can serve as crucial constraints for various high-level tasks [46,47]. Since HDR and LDR images are both 2D images, can we directly use the LDR IQA metrics to evaluate the quality of the HDR images? To answer this question, it is crucial to identify the differences between HDR and LDR images. In HDR images, pixel values are linearly related to the natural luminance. On the contrary, during the capture process of LDR images (as shown in Figure 2), the camera sensors first involve clipping operation to restrict the maximum value of the image. Consequently, the nonlinear mapping process compresses the dynamic range of the image using the Camera Response Functions (CRFs). Finally, the quantization step aims to alleviate the storage burden on cameras. In Figure 3, we provide a more intuitive way to illustrate the statistical differences between the HDR and LDR images. Herein, it is obvious that the HDR image has a wider dynamic range. Furthermore, there are over-exposed regions where many pixel values are 255 in the LDR image. However, there are no such regions in

the HDR image. Therefore, directly applying IQA metrics specifically designed for LDR images to assess the quality of HDR images is unreasonable due to the different data distributions between HDR and LDR images. In Zerman *et al.* [25] and Liu *et al.* [27]'s works, extensive experimental results also demonstrate a significant decline in the performance of the LDR IQA algorithms when they are directly applied to HDR images without any perceptual mapping functions.

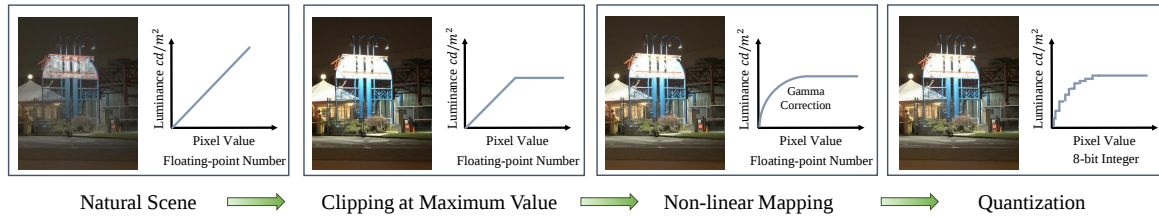


Figure 2. The production process of the LDR image in the sensor.

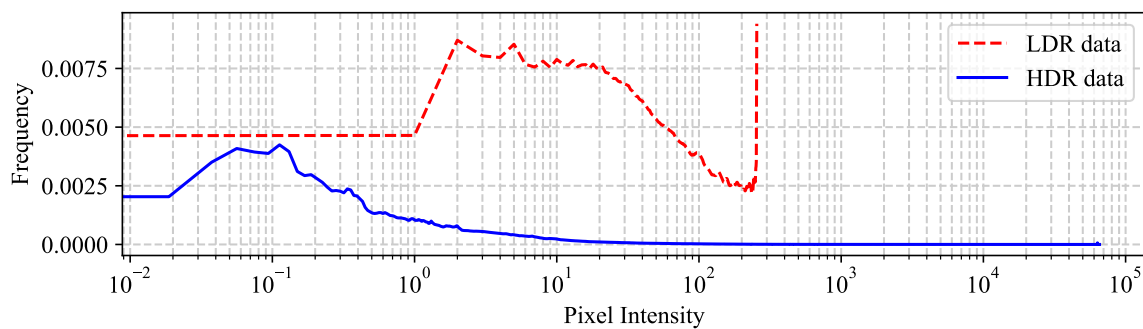


Figure 3. Average histogram of the HDR and LDR images in the SIHDR [48] database.

Consequently, how to effectively evaluate the quality of the HDR images remains a challenging task. Existing HDR IQA methods can be roughly divided into three categories according to the design strategies: 1) HVS-modeling-based HDR IQA methods: directly evaluating the image quality by modeling how the human eye perceives and processes the light signals from natural scenes; 2) Existing IQA metrics based HDR IQA methods: such algorithms aim to make full use of existing HDR or LDR IQA metrics to effectively assess the quality of HDR images; 3) PU-encoding-based HDR IQA methods: assessing the quality of HDR images specifically within the PU domain. More details of existing HDR IQA methods are summarized in Figure 4, Table 2 and described in Sections 3.1, 3.3, and 3.2, respectively.

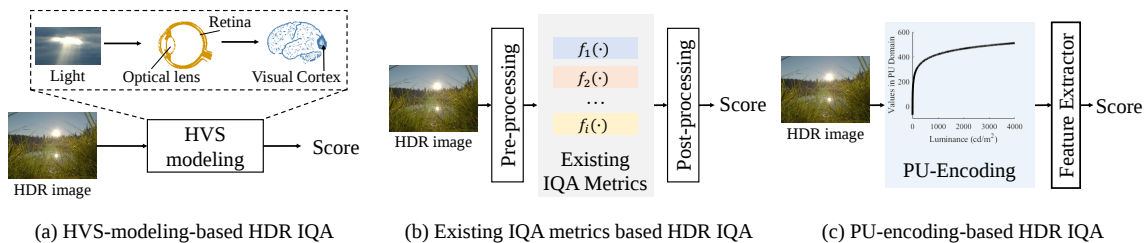


Figure 4. Overview of objective HDR IQA methods. The HDR images presented in this paper are tone-mapped for visualization by Gamma correction.

3.1. HVS-Modeling-Based HDR IQA Methods

HDR-VDP family is a set of widely used and efficient HDR IQA algorithms, where the first version, HDR-VDP-1 [8], is proposed by Mantiuk *et al.*, followed by the improved HDR-VDP-2 [9], HDR-VDP-2.2 [36], and the latest HDR-VDP-3 [10]. The basic idea of the HDR-VDP family is to model

the HVS. Inspired by the Visual Difference Predictor (VDP) [49], HDR-VDP-1 first models the HVS within three stages: 1) modeling the light scattering using the Optical Transfer Function (OTF) [50]; 2) modeling the nonlinear response of the photoreceptors by Just Noticeable Difference (JND); 3) remove the less sensitive frequency according to the Contrast Sensitive Function (CSF). Subsequently, the extracted visual features are decomposed and then fed into a pooling module to calculate the visual difference map. Herein, no single-quality score is provided in HDR-VDP-1. In HDR-VDP-2, the multi-band pooling strategy is proposed to obtain the quality score of the distorted HDR image. Furthermore, the response of the photoreceptors is modeled by sensitive functions [51] instead of the JND. In the pooling stage, the weighting parameters are optimized using the LIVE database [52] due to the lack of HDR IQA databases. Narwaria *et al.* improve the performance of HDR-VDP-2 by optimizing the parameters using two HDR databases [5,23] in HDR-VDP-2.2. Recently, the HDR-VDP-3 metric is proposed, where the impact of different ages on visual quality is considered, together with the modeling of the local luminance. Additionally, the UPIQ database is used for parameter calibration to enable the quality evaluation of both HDR and LDR images. Despite significant success in HDR IQA, the HDR-VDP family still presents aspects worthy of improvement. Firstly, the parameters involved in simulating the HVS and the pooling stage necessitate intricate calibration processes [7]. Secondly, based on the experimental results in [5], HDR-VDP-2 appears to over-penalize the distortions for HDR images with higher quality.

Aydin *et al.* [53] propose a Dynamic Range Independent Metric (DRIM) to measure the visual difference between two images with arbitrary dynamic range. Herein, they first model the HVS to extract the contrast features. Afterward, the cortex transform [54] is applied to obtain three distortion types separately, which are later used to generate the final visual difference map. Similar to HDR-VDP-1, this method tends to predict the probability of the difference between two images, yet it lacks the quantitative evaluation result, such as the MOS value. Given the same distorted image, Aydin *et al.* [34] found that the perceived distortion is more pronounced when viewed on an HDR display compared to an LDR display. Motivated by the above observation, Kottayil *et al.* [55] propose a Deep Neural Network (DNN) based blind IQA for HDR images. The network is composed of an E-Net for error prediction and a P-Net to emulate the perceptual resilience of the HVS, thereby quantifying the visibility of discrepancies within image blocks. The network is trained on the database that combines 5 HDR databases [5–7,23,25], where the MOS values are aligned by the Iterated Nested Least Square Algorithm (INLSA) [56]. Despite employing 5-fold cross-validation, the fact remains that only 552 distorted images are utilized for training in each fold, which is notably insufficient for a DNN model.

3.2. PU-Encoding-Based HDR IQA Methods

PU encoding introduced by Aydin *et al.* [34] aims to convert HDR images into a perceptual domain, enabling the HDR image quality evaluation by LDR IQA metrics, which has been proven to be an effective mapping method [25]. Therefore, existing PU-encoding-based HDR IQA methods apply PU encoding first, followed by the quality evaluation in the PU domain. HDR Video Quality Metric (HDR-VQM) [11] is the first HDR video quality assessment metric, and its efficacy in evaluating the quality of HDR images has been substantiated through subjective experiments [12,25,26]. In this metric, the PU encoding is first applied to obtain the perceived luminance map. Subsequently, the perceptual error is calculated by the log-Gabor [57] filter. The final quality score of the video is predicted by a spatial and temporal pooling strategy. The network proposed by Jia *et al.* [58] design a Convolutional Neural Network (CNN) for blind HDR IQA. In the experiment, the network trained on PU-encoded HDR images shows significant improvement compared with the network trained on original HDR images. The PU-PieAPP proposed by Mikhailiuk *et al.* [26] combines the PU encoding with the PieAPP [59] for the quality evaluation of both HDR and LDR images. The network is trained on the UPIQ database [26], including 80 HDR images and 3779 LDR images. Liu *et al.* [12] introduce the Local-Global Frequency feature-based Model (LGFM). Herein, the odd log-Gabor filter is applied to obtain the local frequency feature, while the Butterworth filter is employed to simulate the CSF to

generate the global frequency feature. The two feature maps undergo separate pooling processes and are subsequently integrated to yield the perceptual quality score.

3.3. Existing IQA Metrics Based HDR IQA Methods

Currently, there is a limited number of HDR IQA algorithms, whereas there are numerous efficient LDR IQA methods. Therefore, it is common practice to directly employ existing LDR IQA methods for quality evaluation after mapping HDR images to the perceptual domain. In the early stage, Aydin *et al.* [34] introduce the PU encoding to convert HDR images into the perceptual domain. Recently, Mantiuk and Azimi propose the PU21 [60] to model the new perceptual uniform space based on a new CSF [61]. Additionally, PQ [35] and Hybrid-Log-Gamma (HLG) [62] are two commonly used transfer functions in codecs, aiming to provide the optimal visualization results of the HDR content on the monitors. Note that PQ and HLG are not dedicated designed for HDR IQA, but they can also be adopted as the perceptual mapping functions. Although these mapping functions can effectively convert the HDR image into the perceptual domain, they are usually non-differentiable, making them unsuitable for the optimization of the learning-based model. Currently, there are four widely used differentiable mapping functions, including the logarithm function, Gamma correction, hyperbolic tangent function (tanh), and μ -law, which can be used to optimize the learning-based models [63–65]. To be specific, the logarithm function and Gamma correction model the nonlinear perception of light in the HVS, which are derived from the Weber-Fechner law and Stevens’ power law [66], respectively. Tanh is an extensively employed nonlinear function that can compress the dynamic range of the image into $[-1, 1]$. The effectiveness of the μ -law algorithm has been demonstrated in various HDR-related tasks [67,68] for dynamic range compression, which is a differentiable companding algorithm originally designed for audio signals.

Table 2. Overview of existing HDR IQA methods.

Method	Year	Classification	Major Category
HDR-VDP-1 [8]	2005	Full-Reference	HVS-modeling-based HDR IQA methods
DRIM [53]	2008	Full-Reference	HVS-modeling-based HDR IQA methods
HDR-VDP-2 [9]	2011	Full-Reference	HVS-modeling-based HDR IQA methods
HDR-VDP-2.2 [36]	2015	Full-Reference	HVS-modeling-based HDR IQA methods
HDR-VQM [11]	2015	Full-Reference	PU-encoding-based HDR IQA methods
Kottayil <i>et al.</i> [55]	2017	No-Reference	HVS-modeling-based HDR IQA methods
Jia <i>et al.</i> [58]	2017	No-Reference	PU-encoding-based HDR IQA methods
HDR-CQM [13]	2018	Full-Reference	Existing IQA metrics based HDR IQA methods
PU-PieAPP [26]	2021	Full-Reference	PU-encoding-based HDR IQA methods
HDR-VDP-3 [10]	2023	Full-Reference	HVS-modeling-based HDR IQA methods
LGFM [12]	2023	Full-Reference	PU-encoding-based HDR IQA methods
Cao <i>et al.</i> [14]	2024	Full-Reference	Existing IQA metrics based HDR IQA methods

In addition, Choudhury and Daly propose to evaluate the quality of HDR images by combining the results of various HDR IQA metrics (HDR-CQM) [13]. To be specific, they first select six HDR IQA metrics with top performance according to the experimental results in [69]. Subsequently, the Sequential Floating Forward Selection (SFFS) [70] algorithm is applied to find the combination of metrics with the best performance. Afterward, the Support Vector Machine (SVM) is trained to predict the final quality score. Cao *et al.* [14] recently design an inverse display model to convert the HDR image into the LDR domain to improve the performance of existing LDR metrics. Specifically, a set of LDR images with different exposures is generated from the proposed inverse display model. The LDR image stacks are used to obtain the local exposure weights, which are further applied for local pooling of the difference map generated from the LDR metric to calculate the final quality score of the HDR image.

4. Discussion and Conclusions

This paper provides a comprehensive overview of existing subjective HDR IQA databases and objective HDR IQA metrics. While these researches represent significant advancements, HDR IQA is still a challenging task, which is evolving alongside the generation of HDR content and the development of compression schemes. Therefore, this section further discusses the limitations of existing methods, along with the potential future work, aiming to provide a comprehensive understanding of the challenges and opportunities in the field of HDR IQA.

For the construction of HDR IQA subjective databases, we can observe three main limitations. 1) *There is a limited number of both reference and distorted HDR images in existing databases.* The content diversity of a database is crucial, as it not only affects the generalizability of the objective IQA algorithm evaluated on it but also impacts the robustness of the DNNs. The DNNs, owing to their powerful feature extraction capabilities, have been applied in numerous IQA models. However, this necessitates a large amount of training data for support. Unfortunately, the current HDR IQA databases are insufficient to train generalized deep learning-based HDR IQA models. 2) *The databases currently focus primarily on limited compression distortions.* As compression technologies evolve, commonly used compression techniques for HDR images have transitioned from JPEG and JPEG2K to JPEG-Xt and Versatile Video Coding (VVC) [71]. Consequently, research on the effects of new compression distortion types on the human visual system remains insufficient. Moreover, there is a lack of research on other types of distortions, including motion distortions that may occur during the capture process, noise, and contrast distortions that may arise during the display process. 3) *All subjective experiments in existing HDR databases are conducted based on professional HDR monitors.* While these monitors offer high brightness levels, they are relatively costly and have limited application scopes. In contrast, currently widely used commercial HDR displays, although unable to achieve the brightness levels of professional monitors, still provide satisfactory visual experiences. Therefore, there is a need for subjective experiments on commercial displays to more accurately cater to the needs of a wider range of application scenarios. Based on the insights above, generating a larger database, exploring the effects of various distortion types, as well as conducting subjective experiments on commercial HDR displays can be the potential research directions in the subjective HDR IQA database construction.

Despite the proliferation of HDR content, there are relatively few algorithms specifically designed to evaluate the quality of HDR images. Accurate IQA models are the foundation of various tasks such as compression or generation. It is still a valuable direction to design robust, reliable, and perceptually aligned mapping methods that can map the brightness and color of the natural world to the perceptual domain that can be better aligned with HVS. The current challenge in evaluating HDR image quality lies in effectively modeling the HVS. Existing methods typically fall into mapping HDR images to the perceptual domain. HVS-modeling-based approaches rely on modeling the HVS based on CSF and luminance masking. However, these methods require numerous parameters obtained from subjective experiments, making the modeling process complex and unsuitable for the optimization of DNNs. On the other hand, PU-encoding-based methods employ a non-linear function modeled

on CSF, facilitating the direct utilization of existing LDR metrics or the design of new IQA metrics. Furthermore, Cao et al.'s research also involves the conversion of HDR images into the LDR domain. Interestingly, evaluation results from [72] indicate that employing PQ or HLG with a current LDR metric surpasses the performance of existing HDR-specific methods, underscoring the significance of perceptual mapping techniques in HDR IQA. Another noteworthy aspect is that the HDR IQA methods are evaluated on existing databases. Consequently, the assessment of these algorithms' effectiveness on other types of distortion is lacking, leading to limitations in current findings. In the future, the objective HDR IQA may involve efficient modeling of the HVS and further refinement of perceptual mapping techniques. These efforts aim to enhance the accuracy and robustness of HDR IQA for improved user experiences across different applications and devices.

Author Contributions: Conceptualization, Y.L. and Y.T.; investigation, Y.L.; resources, Y.T.; data curation, Y.L.; writing—original draft preparation, Y.L. and Y.T.; writing—review and editing, S.W. and X.Z.; visualization, Y.L. and Y.T.; supervision, S.K. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HDR	High Dynamic Range
LDR	Low Dynamic Range
HVS	Human Visual System
IQA	Image Quality Assessment
TMO	Tone Mapping Operator
DSIS	Double-Stimulus Impairment Scale
ACR-HR	Absolute Category Rating with Hidden Reference
CSF	Contrast Sensitive Function
PU	Perceptual Uniform
PQ	Perceptual Quantizer
HLG	Hybrid-Log-Gamma
MOS	Mean Opinion Score
VDP	Visual Difference Predictor
OTF	Optical Transfer Function
DNN	Deep Neural Network

References

1. Metzler, C.A.; Ikoma, H.; Peng, Y.; Wetzstein, G. Deep optics for single-shot high-dynamic-range imaging. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1375–1385.
2. Shopovska, I.; Stojkovic, A.; Aelterman, J.; Van Hamme, D.; Philips, W. High-Dynamic-Range Tone Mapping in Intelligent Automotive Systems. *Sensors* **2023**, *23*, 5767.
3. Sánchez, D.; Gómez, S.; Mauricio, J.; Freixas, L.; Sanuy, A.; Guixé, G.; López, A.; Manera, R.; Marín, J.; Pérez, J.M.; others. HRFlexToT: a high dynamic range ASIC for time-of-flight positron emission tomography. *IEEE Transactions on Radiation and Plasma Medical Sciences* **2021**, *6*, 51–67.
4. Wang, Z.; Chen, W.; Xing, J.; Zhang, X.; Tian, H.; Tang, H.; Bi, P.; Li, G.; Zhang, F. Extracting vegetation information from high dynamic range images with shadows: A comparison between deep learning and threshold methods. *Computers and Electronics in Agriculture* **2023**, *208*, 107805.
5. Narwaria, M.; Da Silva, M.P.; Le Callet, P.; Pepion, R. Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality. *Optical Engineering* **2013**, *52*, 102008–102008.

6. Korshunov, P.; Hanhart, P.; Richter, T.; Artusi, A.; Mantiuk, R.; Ebrahimi, T. Subjective quality assessment database of HDR images compressed with JPEG XT. 2015 seventh international workshop on quality of multimedia experience (QoMEX). IEEE, 2015, pp. 1–6.
7. Valenzise, G.; De Simone, F.; Lauga, P.; Dufaux, F. Performance evaluation of objective quality metrics for HDR image compression. Applications of Digital Image Processing XXXVII. SPIE, 2014, Vol. 9217, pp. 78–87.
8. Mantiuk, R.; Daly, S.J.; Myszkowski, K.; Seidel, H.P. Predicting visible differences in high dynamic range images: model and its calibration. Human Vision and Electronic Imaging X. SPIE, 2005, Vol. 5666, pp. 204–214.
9. Mantiuk, R.; Kim, K.J.; Rempel, A.G.; Heidrich, W. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on graphics (TOG)* **2011**, 30, 1–14.
10. Mantiuk, R.K.; Hammou, D.; Hanji, P. HDR-VDP-3: A multi-metric for predicting image differences, quality and contrast distortions in high dynamic range and regular content. *arXiv preprint arXiv:2304.13625* **2023**.
11. Narwaria, M.; Da Silva, M.P.; Le Callet, P. HDR-VQM: An objective quality measure for high dynamic range video. *Signal Processing: Image Communication* **2015**, 35, 46–60.
12. Liu, Y.; Ni, Z.; Wang, S.; Wang, H.; Kwong, S. High dynamic range image quality assessment based on frequency disparity. *IEEE Transactions on Circuits and Systems for Video Technology* **2023**.
13. Choudhury, A.; Daly, S. HDR image quality assessment using machine-learning based combination of quality metrics. 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2018, pp. 91–95.
14. Cao, P.; Mantiuk, R.K.; Ma, K. Perceptual Assessment and Optimization of HDR Image Rendering. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 22433–22443.
15. Rousselot, M.; Le Meur, O.; Cozot, R.; Ducloux, X. Quality assessment of HDR/WCG images using HDR uniform color spaces. *Journal of Imaging* **2019**, 5, 18.
16. Ward, G.; Simmons, M. JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG. In *ACM SIGGRAPH 2006 Courses*; 2006; pp. 3–es.
17. Sugiyama, N.; Kaida, H.; Xue, X.; Jinno, T.; Adami, N.; Okuda, M. HDR image compression using optimized tone mapping model. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2009, pp. 1001–1004.
18. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **2004**, 13, 600–612.
19. Kuang, J.; Johnson, G.M.; Fairchild, M.D. iCAM06: A refined image appearance model for HDR image rendering. *Journal of Visual Communication and Image Representation* **2007**, 18, 406–414.
20. ITU-T RECOMMENDATION, P. Subjective video quality assessment methods for multimedia applications **2008**.
21. <https://hdr.sim2.it/>.
22. Narwaria, M.; Lin, W.; McLoughlin, I.V.; Emmanuel, S.; Chia, L.T. Fourier transform-based scalable image quality measure. *IEEE Transactions on Image Processing* **2012**, 21, 3364–3377.
23. Narwaria, M.; Da Silva, M.P.; Le Callet, P.; Pépion, R. Impact of tone mapping in high dynamic range image compression. VPQM, 2014, pp. pp–1.
24. Mantiuk, R.; Myszkowski, K.; Seidel, H.P. A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)* **2006**, 3, 286–308.
25. Zerman, E.; Valenzise, G.; Dufaux, F. An extensive performance evaluation of full-reference HDR image quality metrics. *Quality and User Experience* **2017**, 2, 1–16.
26. Mikhailiuk, A.; Pérez-Ortiz, M.; Yue, D.; Suen, W.; Mantiuk, R.K. Consolidated dataset and metrics for high-dynamic-range image quality. *IEEE Transactions on Multimedia* **2021**, 24, 2125–2138.
27. Liu, Y.; Ni, Z.; Chen, P.; Wang, S.; Kwong, S. HDRC: a subjective quality assessment database for compressed high dynamic range image. *International Journal of Machine Learning and Cybernetics* **2024**, pp. 1–16.
28. Reinhard, E.; Stark, M.; Shirley, P.; Ferwerda, J. Photographic tone reproduction for digital images. *ACM Transactions on Graphics* **2002**, 21, 267–276.

29. Durand, F.; Dorsey, J. Fast bilateral filtering for the display of high-dynamic-range images. *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, 2002, pp. 257–266.
30. Ashikhmin, M. A tone mapping algorithm for high contrast images. *Proceedings of the 13th Eurographics workshop on Rendering*, 2002, pp. 145–156.
31. Richter, T. On the standardization of the JPEG XT image compression. *2013 Picture Coding Symposium (PCS)*. IEEE, 2013, pp. 37–40.
32. Series, B. Methodology for the subjective assessment of the quality of television pictures. *Recommendation ITU-R BT* **2012**, 500.
33. Mai, Z.; Mansour, H.; Mantiuk, R.; Nasiopoulos, P.; Ward, R.; Heidrich, W. Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image and Video Compression. *IEEE Transactions on Image Processing* **2011**, 20, 1558–1571.
34. Aydın, T.O.; Mantiuk, R.; Seidel, H.P. Extending quality metrics to full luminance range images. *Human vision and electronic imaging xiii*. SPIE, 2008, Vol. 6806, pp. 109–118.
35. Miller, S.; Nezamabadi, M.; Daly, S. Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal* **2013**, 122, 52–59.
36. Narwaria, M.; Mantiuk, R.K.; Da Silva, M.P.; Le Callet, P. HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images. *Journal of Electronic Imaging* **2015**, 24, 010501–010501.
37. Perez-Ortiz, M.; Mikhailiuk, A.; Zerman, E.; Hulusic, V.; Valenzise, G.; Mantiuk, R.K. From pairwise comparisons and rating to a unified quality scale. *IEEE Transactions on Image Processing* **2019**, 29, 1139–1151.
38. Pan, Z.; Yuan, F.; Lei, J.; Fang, Y.; Shao, X.; Kwong, S. VCRNet: Visual compensation restoration network for no-reference image quality assessment. *IEEE Transactions on Image Processing* **2022**, 31, 1613–1627.
39. Ding, K.; Ma, K.; Wang, S.; Simoncelli, E.P. Image quality assessment: Unifying structure and texture similarity. *IEEE transactions on pattern analysis and machine intelligence* **2020**, 44, 2567–2581.
40. Tian, Y.; Wang, S.; Chen, B.; Kwong, S. Causal Representation Learning for GAN-Generated Face Image Quality Assessment. *IEEE Transactions on Circuits and Systems for Video Technology* **2024**.
41. Ni, Z.; Liu, Y.; Ding, K.; Yang, W.; Wang, H.; Wang, S. Opinion-Unaware Blind Image Quality Assessment using Multi-Scale Deep Feature Statistics. *IEEE Transactions on Multimedia* **2024**.
42. Chen, B.; Zhu, H.; Zhu, L.; Wang, S.; Kwong, S. Deep Feature Statistics Mapping for Generalized Screen Content Image Quality Assessment. *IEEE Transactions on Image Processing* **2024**.
43. Li, Y.; Chen, B.; Chen, B.; Wang, M.; Wang, S.; Lin, W. Perceptual quality assessment of face video compression: A benchmark and an effective method. *IEEE Transactions on Multimedia* **2024**.
44. Zhu, H.; Wu, H.; Li, Y.; Zhang, Z.; Chen, B.; Zhu, L.; Fang, Y.; Zhai, G.; Lin, W.; Wang, S. Adaptive Image Quality Assessment via Teaching Large Multimodal Model to Compare. *arXiv preprint arXiv:2405.19298* **2024**.
45. Tian, Y.; Chen, B.; Wang, S.; Kwong, S. Towards Thousands to One Reference: Can We Trust the Reference Image for Quality Assessment? *IEEE Transactions on Multimedia* **2023**.
46. Lei, J.; Li, D.; Pan, Z.; Sun, Z.; Kwong, S.; Hou, C. Fast intra prediction based on content property analysis for low complexity HEVC-based screen content coding. *IEEE Transactions on Broadcasting* **2016**, 63, 48–58.
47. Hong, Y.; Kwong, S.; Chang, Y.; Ren, Q. Consensus unsupervised feature ranking from multiple views. *Pattern Recognition Letters* **2008**, 29, 595–602.
48. Hanji, P.; Mantiuk, R.; Eilertsen, G.; Hajisharif, S.; Unger, J. Comparison of single image HDR reconstruction methods—the caveats of quality assessment. *ACM SIGGRAPH 2022 conference proceedings*, 2022, pp. 1–8.
49. Daly, S.J. Visible differences predictor: an algorithm for the assessment of image fidelity. *Human Vision, Visual Processing, and Digital Display III*. SPIE, 1992, Vol. 1666, pp. 2–15.
50. Deeley, R.J.; Drasdo, N.; Charman, W.N. A simple parametric model of the human ocular modulation transfer function. *Ophthalmic and Physiological Optics* **1991**, 11, 91–93.
51. Stockman, A.; Sharpe, L.T. The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision research* **2000**, 40, 1711–1737.
52. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing* **2006**, 15, 3440–3451.

53. Aydin, T.O.; Mantiuk, R.; Myszkowski, K.; Seidel, H.P. Dynamic range independent image quality assessment. *ACM Transactions on Graphics (TOG)* **2008**, *27*, 1–10.
54. WATSON, A.; others. The cortex transform- Rapid computation of simulated neural images. *Computer vision, graphics, and image processing* **1987**, *39*, 311–327.
55. Kottayil, N.K.; Valenzise, G.; Dufaux, F.; Cheng, I. Blind quality estimation by disentangling perceptual and noisy features in high dynamic range images. *IEEE Transactions on Image Processing* **2017**, *27*, 1512–1525.
56. Pinson, M.H.; Wolf, S. An objective method for combining multiple subjective data sets. *Visual Communications and Image Processing 2003. SPIE, 2003, Vol. 5150, pp. 583–592.*
57. Field, D.J. Relations between the statistics of natural images and the response properties of cortical cells. *Josa a* **1987**, *4*, 2379–2394.
58. Jia, S.; Zhang, Y.; Agrafiotis, D.; Bull, D. Blind high dynamic range image quality assessment using deep learning. *2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017, pp. 765–769.*
59. Prashnani, E.; Cai, H.; Mostofi, Y.; Sen, P. Pieapp: Perceptual image-error assessment through pairwise preference. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1808–1817.*
60. Azimi, M.; others. PU21: A novel perceptually uniform encoding for adapting existing quality metrics for HDR. *2021 Picture Coding Symposium (PCS). IEEE, 2021, pp. 1–5.*
61. Mantiuk, R.K.; Kim, M.; Ashraf, M.; Xu, Q.; Luo, M.R.; Martinovic, J.; Wuerger, S. Practical Color Contrast Sensitivity Functions for Luminance Levels up to 10000 cd/m². *Color and Imaging Conference. Society for Imaging Science & Technology, 2020, Vol. 28, pp. 1–6.*
62. Borer, T.; Cotton, A. A display-independent high dynamic range television system. *SMPTE Motion Imaging Journal* **2016**, *125*, 50–56.
63. Huang, X.; Zhang, Q.; Feng, Y.; Li, H.; Wang, X.; Wang, Q. Hdr-nerf: High dynamic range neural radiance fields. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 18398–18408.*
64. Chen, X.; Liu, Y.; Zhang, Z.; Qiao, Y.; Dong, C. Hdrunet: Single image hdr reconstruction with denoising and dequantization. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 354–363.*
65. Catley-Chandar, S.; Tanay, T.; Vandroux, L.; Leonardis, A.; Slabaugh, G.; Pérez-Pellitero, E. Flexhdr: Modeling alignment and exposure uncertainties for flexible hdr imaging. *IEEE Transactions on Image Processing* **2022**, *31*, 5923–5935.
66. Stevens, S.S. On the psychophysical law. *Psychological review* **1957**, *64*, 153.
67. Santos, M.S.; Tsang, R.; Khademi Kalantari, N. Single Image HDR Reconstruction Using a CNN with Masked Features and Perceptual Loss. *ACM Transactions on Graphics* **2020**, *39*. doi:10.1145/3386569.3392403.
68. Chen, J.; Yang, Z.; Chan, T.N.; Li, H.; Hou, J.; Chau, L.P. Attention-guided progressive neural texture fusion for high dynamic range image restoration. *IEEE Transactions on Image Processing* **2022**, *31*, 2661–2672.
69. Hanhart, P.; Bernardo, M.V.; Pereira, M.; G. Pinheiro, A.M.; Ebrahimi, T. Benchmarking of objective quality metrics for HDR image quality assessment. *EURASIP Journal on Image and Video Processing* **2015**, *2015*, 1–18.
70. Pudil, P.; Novovičová, J.; Kittler, J. Floating search methods in feature selection. *Pattern recognition letters* **1994**, *15*, 1119–1125.
71. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on circuits and systems for video technology* **2012**, *22*, 1649–1668.
72. Sugito, Y.; Vazquez-Corral, J.; Canham, T.; Bertalmío, M. Image quality evaluation in professional HDR/WCG production questions the need for HDR metrics. *IEEE Transactions on Image Processing* **2022**, *31*, 5163–5177.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.