Reduced reference image and video quality assessments: review of methods

Shahi Dost[1*] , Faryal Saud[2], Maham Shabbir[2], Muhammad Gufran Khan[3], Muhammad Shahid[4] and Benny Lovstrom[5]

*Correspondence:
shahidost@gmail.com
[1] TIB – Leibniz Information
Centre for Science
and Technology, Hannover,
Germany
Full list of author information
is available at the end of the
article

Abstract

With the growing demand for image and video-based applications, the requirements of consistent quality assessment metrics of image and video have increased. Different approaches have been proposed in the literature to estimate the perceptual quality of images and videos. These approaches can be divided into three main categories; full reference (FR), reduced reference (RR) and no-reference (NR). In RR methods, instead of providing the original image or video as a reference, we need to provide certain features (i.e., texture, edges, etc.) of the original image or video for quality assessment. During the last decade, RR-based quality assessment has been a popular research area for a variety of applications such as social media, online games, and video streaming. In this paper, we present review and classification of the latest research work on RR-based image and video quality assessment. We have also summarized different databases used in the field of 2D and 3D image and video quality assessment. This paper would be helpful for specialists and researchers to stay well-informed about recent progress of RR-based image and video quality assessment. The review and classification presented in this paper will also be useful to gain understanding of multimedia quality assessment and state-of-the-art approaches used for the analysis. In addition, it will help the reader select appropriate quality assessment methods and parameters for their respective applications.

Keywords: Multimedia quality assessment, Reduced reference image quality approaches, Image quality parameters, Reduced reference video quality approaches, Video quality parameters

1 Introduction

The demand for image and video quality assessment is growing rapidly in the emerging multimedia applications. During transmission and processing of multimedia,[1] distortions often occur which may cause degradation of the visual quality. The final display of multimedia contents to a human viewer goes through many stages of processing and signal distortions may be introduced by any of these stages. It becomes necessary to measure and quantify the quality of multimedia and to scale the distortions that have been added at various stages.

---

[1] The focus of this paper is on image and video contents only.

The studies in the field of image and video quality assessment aim at the development of metrics that can be used to calculate the quality of multimedia contents. The parameters and assessment metrics used in quality estimation methods play a significant role in a wide range of systems like modern video broadcasting systems including high definition TV, and applications such as image acquisition, displaying, enhancement, restoration, compression, printing, analysis, and watermarking [1].

Since Quality of Experience (QoE) is considered a better estimate of perceptual quality than the Quality of Service (QoS) [2, 3], it is important that the multimedia quality assessment is quantified through human perception. To judge the quality of an image or a video, the most reliable method is undoubtedly the *subjective evaluation* as human observers are the ultimate receivers of the content. A subjective assessment of quality using human observers can provide us a Mean Opinion Score (MOS). This assessment technique has been used previously as a subjective quality assessment metrics for the Image Quality Assessment (IQA) as well as Video Quality Assessment (VQA). There are differences between the approaches used for image and video quality assessment; given the fact that the mechanism of quality distortion are mostly different for images and videos. In this article, we address and present both image and video quality assessment techniques and classify them accordingly. But subjective assessment is time-consuming and expensive approach and cannot be implemented in real-time scenarios. Therefore, *objective methods* of quality assessment are required. A distortion-free and perfect quality image or video can be used as a reference against a distorted signal, which may give us an objective quality assessment. According to the degree of information available for the reference image or video signals, we can classify the quality assessment methods into full reference (FR), reduced reference (RR) and no-reference (NR) categories, as described below:

- *Full reference (FR)* In this category of methods, the reference multimedia contents are fully available for comparison with the received distorted contents in order to evaluate visual quality. However, in most practical applications, the original signal is not available at the client or receiver end. Using FR approach the metrics such as Structural Similarity Image Index (SSIM) [4], Multi-Scale Structural Similarity Index (MS-SSIM) [5], Feature Similarity Index (FSIM) [6], Gradient Magnitude Similarity Deviation (GMSD) [7], and Perceptual Similarity Index (PSIM) [8] have been proposed.
- *Reduced reference (RR)* In RR methods, it is not necessary to have access of the original multimedia contents for quality assessment purpose. Instead, only the characteristic information about pixels, coefficients of certain transformation, or other dominant features of the original image or video are provided. This is a practical approach for real-time scenarios and some examples of this approach include the reduced reference variants of SSIM and MS-SSIM, Reduced Reference Entropic Differencing (RRED) [9], and Spatial Efficient Entropic Differencing (SpEED) [10].
- *No reference (NR)* The quality assessment of image or video in this category is performed blindly on the basis of extracted features from the multimedia content under assessment as there is no reference available. However, NR-based image and video quality evaluation is a challenging task as the extracted features may provide very

limited information. Some examples of this category include BRISQUE [11], DII-VINE [12], Natural Image Quality Evaluator (NIQE) [13], NR-Free Energy-based Robust Metric (NFERM) [14], (COde- book Representation for No-reference Image quality (CORNIA) [15], BPRI [16] and Blind Multiple Pseudo-Reference Images (BMPRI) [17]. An NR image quality measuring metric by using the free-energy-based brain theory and human visual system (HVS)-inspired features are proposed by [18]. Another approached based on free-energy-based distortion metric (FEDM) and structural degradation model has been proposed by Gu et al. [19].

The multimedia quality assessment employing RR methods provides an intermediate approach between FR and NR as it requires only partial information about multimedia contents on the receiver end [20, 21]. In the RR approach, the features extraction process is performed both at the sender and receiver sides. The extracted features and multimedia content are sent over a medium that is assumed to be error-free. Once the features are sent, at the receiver side the same type of features are extracted from the received media, and different techniques are employed to measure the degradation in the perceptual quality. An optimal RR-based quality assessment method must attain good balance between the amount of data produced by RR features and the correctness of multimedia contents quality assessment. A user can have access to a huge amount of information as reference that can lead to the highly accurate estimation of quality of distorted multimedia contents. Obviously, a huge amount of RR features data are transmitted to the target system in this case. In contrast, if less data are sent on the same channel it takes relatively less time to communicate. As a result, it will be convenient to send lesser RR features data but precision of the quality estimation will suffer.

In multimedia world of today, the end-user demands a high-quality image and video viewing experience. Therefore, the quality assessment techniques have become extremely important and have been used in a wide array of applications. For example, *Internet Services Providers* and *network operators* have a strong interest to deliver high-quality services to the end-users. RR methods provide a quality metrics for the satisfaction of the end-user. There are a number of partners involved between service providers and end-users, which need *service level agreement* to guarantee the agreed quality standard to be provided to the end-user. Therefore, in such cases, RR methods can be an appropriate choice for quality monitoring of the live streaming systems [22, 23]. Moreover, there have been significant advances in the area of image and video compression, and various algorithms have been developed to compressed multimedia contents. RR approaches can be used to measure the quality of these multimedia contents after the compression algorithm.

In order to find appropriate method of assessment for a specific RR quality assessment application, one would need to review the related methods. To the best of our knowledge there is no study and comparison available in literature that addresses RR-based quality assessment methods. This article addresses the shortcomings and presents literature review in a classified way for those readers interested in the areas of RR-based quality assessments. In addition, our contribution presents an overview of RR quality assessment metrics with respect to domain-based classification (i.e., pixel, frequency, and bitstream) that helps in selection of the metrics according to the required multimedia

contents (i.e., image, videos, 2D or 3D based). The authors are expecting that this work can be helpful for specialists and researchers, and will provide review and summary of recent progress (development) in the area of RR-based quality assessment methods.

The rest of the article is structured as follows: In Sect. 2, we present state-of-the-art approaches in the field of multimedia quality assessment with a particular focus on RR-based approaches. In Sect. 3, we present the databases used for the development of new quality assessment parameters. In Sect. 4, we describe our proposed classification for RR-based quality assessment approaches in details.

In Sect. 5, we briefly summarize, conclude the quality assessment approaches, and suggest future works.

2 Related work

The approaches presented in [24, 25] are developed upon natural scene statistics which enable quality assessment to deliver reasonable performance in terms of human perception. However, various challenges in quality assessment design lead to different processes for RR quality evaluation under different circumstances. The algorithms for RR IQA and VQA either use relative entropy or entropic difference. For standard and high-resolution video quality assessment, the reader can refer to the classified models presented by [26, 27]. Visual statistics and visual features are used for basic classification of RR approaches which are further classified into the frequency and pixel-based approaches [27].

A review article for IQA is presented in [28]. In this article, the authors analyze the factors which affect both 2D and 3D image quality and provide the quality measurements of distorted images with respect to these factors. They also described the IQA databases and presented experimental results of IQA metrics. They presented the overall IQA approaches, which lack in differentiating between different FR, RR, and NR-based quality methods. They only target quality parameters that developed for 2D and 3D images quality estimation.

In [29], a survey of quality assessment metrics and the applications of these methods are presented. The importance is given to the metrics that depict quality measures from an end-user perspective. The authors in [26] presented various FR and RR approaches, which are divided into (1) point-based metrics (2) natural visual characteristics, (3) and *Human Visual System* (HVS) metrics. Further, natural visual characteristics (NVC) are divided into two sub-categories, i.e., natural visual statistics and natural visual features. The HVS perceptual metrics are also sub-divided into the frequency and pixel-based methods. The authors in [26, 29] presented an overview of image quality metrics of FR and NR.

One of the related review on perceptual image visual quality metrics presented in [30] is systematic, but its focus is only on six image metrics, namely SSIM [4], PSNR [31, 32], IFC [33], MSVD [34], VSNR [35] and VIF [36]. A survey on NR Image Quality Assessment (IQA) based quality assessment approaches are presented in [37]. The paper presents several frequency-based modules including signal decomposition, visual attention, just noticeable distortion and common feature as a whole without distinguishing FR, NR and RR assessment.

A machine learning-based framework, which utilized saliency detection from multimedia contents are developed in [38]. This framework can predict the quality

measurement for two common types of distortion, noise and JPEG compression. In the first phase, the framework predicts the distortion level and removed that distortion. In the second phase, the saliency map is calculated by using saliency detection algorithms, which measure the amount of distortion added in the multimedia contents. This framework is evaluated on Tampere Image Database (TID2013), which shows overall promising results.

Another review of image quality assessment and different challenges in these fields are highlighted by [39]. They reported key properties of visual perception, quality assessment datasets and existing full, no and reduced reference IQA algorithms.

A survey on frequently used subjective image quality assessment database are reported in [40]. They also classified and reviews objective image quality assessment on the basis of applications and the methodologies utilized in the quality measures. At the end they make performance comparison of quality measures for visual signals with evaluations protocols.

By looking at the literature, either the community focused on individual parameter-based RR approaches [30], or presented a few RR approaches with their NR or FR surveys [37]. There is not a single paper, which described all the quality parameters of RR with respect to multimedia content in a comprehensive way. Due to these shortcomings, we present a study of RR-based quality measuring approaches for both image and video quality assessment. We have presented different classifications of RR image and video quality approaches and compared the metrics performance of each class with respect to their approach. This would help the researcher to select the best approach according to their application and select related literature for the development of new methods.

3 Databases for RR quality assessment approaches

To check the appropriateness of developed RR IQA and VQA methods, a number of databases are used for the evaluations. In order to test the performance of these approaches, researchers usually use publically available databases for the evaluation of their developed quality metric. Some of the widely used public databases are described below:

1. LIVE2005:[2] Image and Video Quality Assessment (LIVE) [41] database is widely used for image and video quality assessment. This database is divided into 5 parts on the basis of distortions types that is (i) JPEG compression of 169; (ii) JPEG2000 compressed of 175; (iii) Gaussian blur of 145; (iv) white noise of 145 images, and (v) bit errors in JPEG2000 bitstream of 145 images.
2. TID2008 [42]: This dataset consists of 1700 test images (25 reference images, 17 types of distortions for each reference image, 4 different levels of each type of distortion). Mean Opinion Scores (MOS) for this database have been obtained as a result of more than 800 experiments.
3. TID2013 [43]: This dataset is mainly developed for full reference visual quality assessment. This is the updated version of TID2008 with larger number (3000) of

---

[2] https://live.ece.utexas.edu/research/quality/subjective.htm.

Dost *et al. EURASIP Journal on Image and Video Processing*   (2022) 2022:1

Page 6 of 31

test images obtained from 25 reference images, 24 types of distortions and 5 levels for each distortion.

4. IVL:[3] The IVL [44] database consists of 20 original images of 886x591 pixels. These images are divided into two different contents in terms of low-level features (i.e., frequencies, colors) and higher (i.e., face, buildings, close-up, outdoor, landscape).

5. The IRCCyN/IVC[4][45] database was developed by the Institute de Recherche en Communications ET Cyberntique de Nantes. This database consists of ten original images and 255 distorted images produced by four different processing methods (JPEG, JPEG2000, LAR coding, and blurring).

6. A well-known database is developed by [46, 47] for video quality assessment. For video quality assessment, the well-known database is developed by [46, 47]. This database consists of 20 videos. All these videos are HD YUV 4:2:0 format sequences which are downsampled to a resolution of 768×432 pixels. The frame rate of these videos is 25 or 50 per second. For every video sequence, 18 distorted versions are produced, with different types of distortion including IP-distortion, Wireless distortion, H.264 compression, MPEG-2 compression.

7. Image and video communications (IVC): Some researchers used the IVC test database, which is a set of 12 original images degraded by 5 types of distortions and quality scores [42, 45, 48].

8. Toyoma: Contains subjective assessment-based test data and stimuli generated via processing of 16 reference images using JPEG and JPEG2000 compression [49, 50].

9. EPFL 3D image database [51]: This database consists of stereoscopic images with a resolution of 1920x1080 pixels. Each scene was captured with fluctuating camera distances in the range of 9.5–61.5 cm. The database contains 11 scenes. For each scene, 6 different cameras are used to capture the scene with respect to different distances.

10. 3D Video IQA database:[5] This database consists of 3D stereoscopic videos with a resolution of 1920×1080 pixels and a frame rate of 25 fps. In this database, various indoor and outdoor scenes with a variety of colors, textures, moving objects and depth structures have been captured. This database has 30 videos and the 30×20 score matrix as CSV files.

A number of databases are also analyzed and presented in [52] for in-depth details. They proposed several criteria for the quantitative comparisons of subjective ratings, test conditions and source content, which are used as the basis for correct analyses and discussion.

4 Classification of RR quality assessment methods
In today's Internet environment, RR-based multimedia quality assessment approaches are widely used, because they provide a features-based technique as compared to FR and NR techniques. NR-based methods are blind and do not utilize any original multimedia
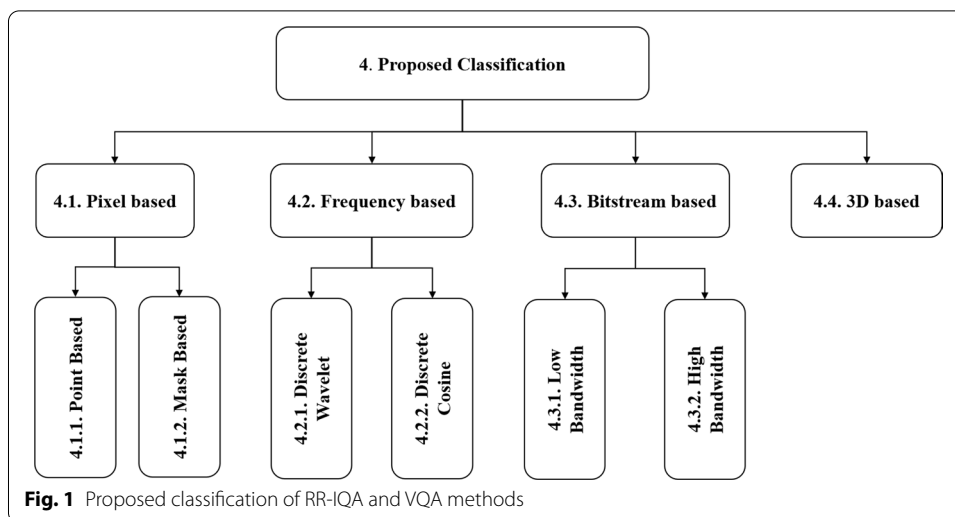
---

[3] http://www.ivl.disco.unimib.it/activities/imagequality/.

[4] http://www2.irccyn.ec-nantes.fr/ivcdb/.

[5] https://mmspg.epfl.ch/downloads/3dvqa/.

Fig. 1 Proposed classification of RR-IQA and VQA methods

contents for quality assessment due to which, we cannot totally rely on their estimation for multimedia quality. On the other hand, FR-based approaches make use of full multimedia contents to measure the quality, which is not used in most of the practical scenarios. The RR-based methods are used in real-time scenarios because it is more reliable as compared to NR and used fewer overhead data as compared to FR-based approaches. These methods use different RR features based on the scenarios, so it is useful to classify these methods into meaningful classes. In literature, there is not a single work that classifies these RR methods into classes and sub-classes on the basis of scenario and multimedia contents.

In most of the scenarios, the quality of the multimedia contents is interpreted in the form of pixels, frequency or bitstream-based operations. The pixel-based operation is performed on the pixels using one or many pixels at a time. Frequency-based methods such as discrete cosine transform (DCT) and wavelet coefficients use frequency transformation of the original image for the quality estimation. Pixel-based methods use the simple process as compared to frequency-based operations, but in some scenarios, frequency-based methods are more efficient because pixel-based methods do not provide sufficient information for quality scores. Bitstream-based methods, on the other hand, make use of stream of bits data which is obtained by channel encoding and decoding. These methods are computationally less intensive due to the whole data decoding for quality estimation.

In our proposed classification, RR methods have been divided into four main categories, i.e., pixel, frequency, bitstream, and 3D multimedia-based[6] methods. From our viewpoint, this classification is very useful with respect to the interpretation of multimedia data. We further classified these classes into sub-classes. The *pixel-based* approaches are divided into the *point* and *mask-based* approaches. Similarly, *frequency-based* methods are divided into *wavelet* and *DCT coefficient*-based methods. *Bitstream-based* methods are divided according to the communication channel into *low* and *high*

---

[6] We focused on both 3D image and video for the RR-based quality estimation methods and classify into one class to differentiate from 2D image and video.

bandwidth-based methods. *3D-multimedia*-based methods are represented in a *single-class* due to the limited approaches developed in this category. The details of classification are presented in the next section with graphical-overview in Fig. 1.

4.1 Pixel-based methods

Many quality assessment methods are available in the literature which directly manipulate the image pixel values for estimating the multimedia contents quality. These methods are divided into *point* (i.e., the individual pixel point of multimedia contents in most of the approaches) and *mask* (i.e., more than one pixel in a small portion of multimedia content) based RR quality methods. Point-based operations work on individuals, while mask-based operations use adjacent pixel values. The classical implementations of these methods are sequential and their complexity is proportional to the number of pixels in an image. Point and mask-based operations are less complex as compared to frequency methods due to the complex operation like discrete Fourier transform (DFT) involved. Many well-known objective quality assessment methods, such as peak signal-to-noise ratio (PSNR) [31] and mean square error (MSE) [53] are pixel-based methods. These methods are used from the beginning of image and video processing and also in the quality assessment field. Due to the pixel and mask-based operations and their importance in the multimedia fields, we have classified these methods into the point and mask-based RR approaches.

4.1.1 Point-based methods

Point-based operations on an image are independent of the values of pixels as a whole and only take into account one pixel at a time and performed operations. In this class of RR methods, every operation takes one pixel and changes its value according to the content of multimedia data (i.e., image or video). These approaches are described below in detail.

The approach discussed in [50] uses three test images and used for features-based PSNR values. The combination of different features constitutes a higher correlation between quality assessment for image or video content. Those images or videos, which are decided for subjective quality score (i.e., the score valued varies between 0 and 1) depends on the contents and their corresponding MOS. The human visual system understands and also predicts the main visual information of image and video. The insight quality of image and video relies on the information of fidelities (i.e., the natural scene statistics and the notion of image/video contents extracted by the human visual system).

In [54], a novel RR-IQA-based approach which used multimedia fidelity of image visual information for features to measure the quality of the image. Their architecture which used full and reduced reference distorted image is shown in Fig. 2.

An orientation selectivity (OS)-based RR-IQA proposed by [55] used the visual contents of extracted OS as RR features for IQA. The mechanism first used to analyze the two nearby pixels similarity and then the orientation similarities of the local neighborhood pixels. With the help of *orientation selectivity visual pattern*, the visual features of an image are measured and plotted in the OS histogram map. The two histograms mapping by reference and the distorted image is measured for quality

Fig. 2 Pixel-based RR-IQA methods with respect to original (reference) and distorted (test) image [54]
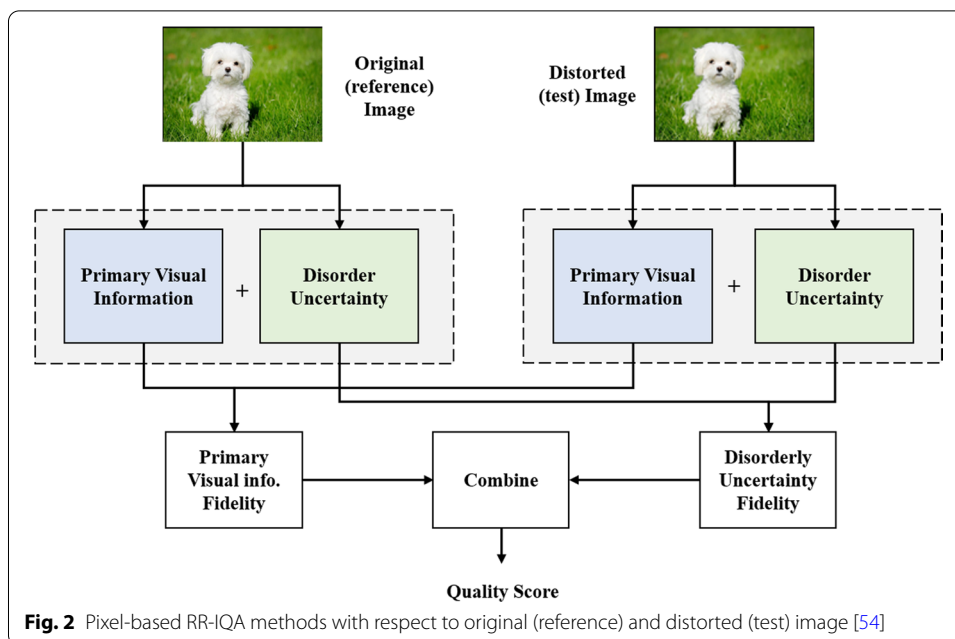
estimation and calculated the changes between the two histograms. The changes from histograms will estimate the quality metrics of the image. The more in changes in a base histogram the greater the quality loses by the original image.

The work presented in [56] is based on the phase congruency (PC) changes between the original and distorted image. The model works in three stages. In the start stage, the fractal dimensions [57] of reference and distorted images are calculated on the PC as features of the image. In the later stage, the image features are characterized by spatial distribution features. In the final stage, the image features are gathered as the quality score using the distance measure. On the basis of these three stages, the quality of the image is measured.

A unique RR-IQA method proposed in [21, 58] is based on exploiting the spatial and temporal information loss and statistical features based on an inter-frame histogram. A proposed *Energy Variation Descriptor* (EVD) measures the energy change in a frame that is caused during the quantization process in spatial domain. EVD also has the ability to depict the texture masking attribute of HVS. The Generalized Gaussian Density (GGD) function is used to capture the inter-frame statistical distribution in the temporal perspective. The City-Block Distance (CBD) works by calculating the histogram differences between video sequences. A proficient RR VQA is developed by merging spatio-temporal features based on EVD and CBD. The proposed method outperforms the FR VQA and RR VQA in subjective evaluations which implies a more precise depiction of HVS. Only a small number of RR features are taken out for expressing the original frame data.

A systematic summary of point-based RR perceptual quality assessment methods is presented in Table 1, with details. Point-based approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance.

In Table 1, the method in [50] uses point operations for the calculation of PSNR which is in turn used for the estimation of quality scores. The authors used both the JPEG and JPEG2000 compression for calculating the distortion made by compression and find their technique exploiting RR-based approach.

4.1.2 Mask-based methods

Mask-based operations on an image take into account a sub-image or a portion of an image for each operation. Each portion contributes to the final output of the operation. Mask-based RR perceptual quality assessment approaches are presented below in detail.

Different types of features, such as linear structure orientation, length, width, maximum magnitude, contrast, and local mean value can be used for assessing the quality of the multimedia content to find the best in terms of quality assessment. The approach presented in [50] has discussed testing of these features for quality assessment. The approach also discussed how the information of these features can be combined to get higher performance. The structural information of image and video frames are very sensitive to noise and image distortion [59]. This structural information can be used for RR quality assessment. The perceptual structural information of image and video are used by [59, 60] for RR image/video features for quality estimation. They used JPEG and JPEG2000 images of structural information for their proposed approach testing.

Visual features information can be used as a measure for video or image quality assessment. The method proposed in [50] is based on the directional edge projections. The edge information is obtained using Sobel filters. Once the image is decompressed or received over the network channel, the edge profiles of original and distorted media are checked to find the quality differences. Statistical features from a multi-scale orientation of image with divisive normalization can be used for RR-IQA [61], the methods used by [62] is based on the structural similarity index (SSIM), which can also be used in the majority of FR image quality estimation measurement.

The relation between depth map quality and overall quality of light field plays an important role to study the quality of the distorted image. RR IQA method based on the depth map and overall quality of light field is presented in [63]. They measured the distortion in the depth map of distorted images by utilizing their own developed dataset for image quality measurement prediction.

The image-statistic is based on image gradient magnitude and Weibull distribution using scale and shape, which can play an important role in the assessment of RR image features [53]. An approach based on scale and shape with image gradient magnitude and Weibull distribution is proposed in [53]. They used the strongest component map in terms of scale-space as RR features for image quality. Singular value decomposition (SVD) method discussed in [64] can be used for the quality estimation. SVD was previously used for image luminance information, but it is also used for image feature extraction. SVD algorithm reduces a high amount of data while retaining high accuracy. The novel SVD approaches employing the multi-scale structural similarity index (MS-SSIM) for quality prediction, which can be used.

The author in [65] proposed an approach, which exploit and integrate analysis of space–time slices with frame-based image quality measurement for video quality prediction. They first arrange the test and reference video sequence into a space–time slice

Fig. 3 The two-layer overview of RR quality assessment system based on mask-based method [67]
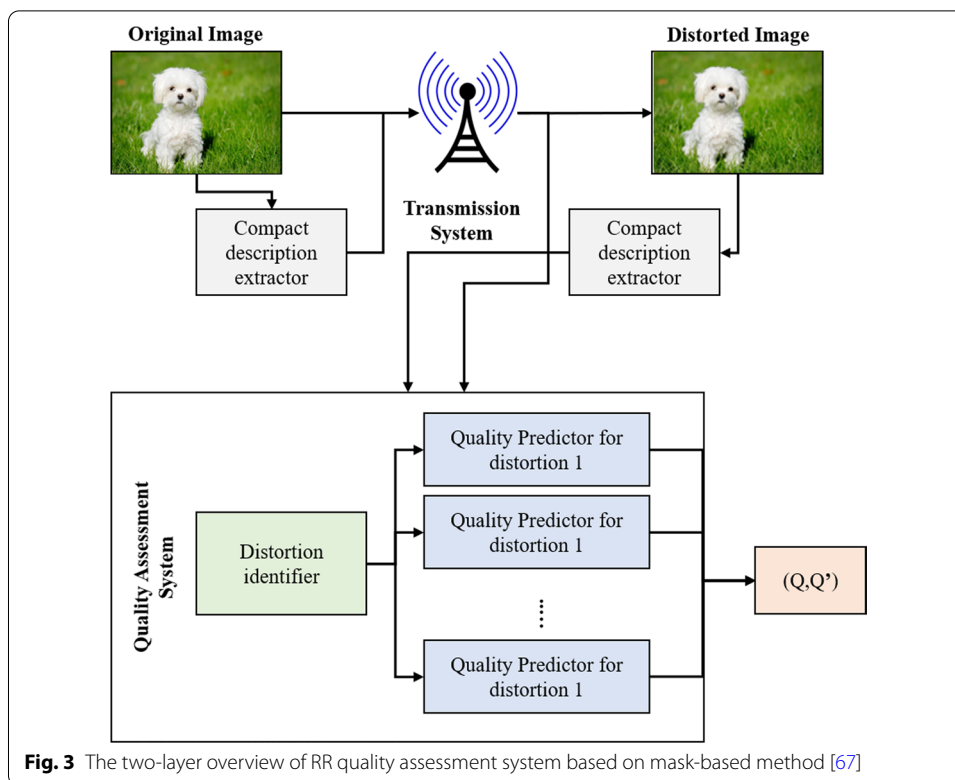
representation. Then they compared a collection of distortion-aware maps with each reference-test video to make pair to distortion measurement.

Fast Johnson–Lindenstrauss transform (FJLT)-based image hashing technique for RR approach provides the low data-rate features of multimedia content for reference and accurately estimates the quality degraded by JPEG and JPEG2000 compression [66]. This technique is robust against many types of distortions including compression. The quantity of FJLT hashing features for RR-based quality assessment technique is small to fulfill the requirement of low data. The FJLT approach is based on three steps: (i) random sampling, (ii) dimension reduction, and (iii) weights incorporation. In random sampling, the image is first converted to grayscale and then 'N' (*i.e., N number of images*) sub-images are selected using the secret key. These features matrix are mapped into lower dimensions using FJLT with minor distortions, and then weights are assigned to hash features randomly. This final information can be used as the RR feature to estimate the quality of multimedia contents degree by distortion-like compression in that particular case.

Two-layer approaches proposed by [67–69] use color correlogram for analyzing the variations in color images. The first layer processes the image and finds the type of distortion in the image. The second layer identifies and predicts the kind of degradation. The color correlogram (i.e., ACF-Auto Correlation Function) finds the alteration in the distribution of color image and two-layer system used for RR image quality estimation for the reference image and finds the quality of the degraded image. The approach is presented in Fig. 3.

A new RR-based system is presented in [66] that uses less than 10 kbps RR features and still achieves high subjective quality. This system is based on similar feature

Dost *et al. EURASIP Journal on Image and Video Processing*   (2022) 2022:1

Page 12 of 31

Table 1 Summary of point and mask-based RR perceptual quality assessment metrics values with respect to the distortion types (i.e., JPEG, JPEG2000)

Domain	Multimedia	Resolution	Distortion type	Quality parameter	Parameter value
Point	Image	512 × 512	JPEG,	PSNR [50]	0.918,
		512 × 512	**JPEG2000**	**PSNR [50]**	**0.924**
		512 × 512	JPEG,	PSNR [54]	0.532
		512 × 512	JPEG2000	PSNR [54]	0.553
		512 × 512	JPEG	PSNR [55]	0.732
		482 × 512	JPEG2000	PSNR [57]	0.872
		512 × 512	JPG	PSNR [61]	0.902
		512 × 512	JPEG,	SSIM [62]	0.773
		512 × 512	JPEG2000	SSIM [62]	0.791
	Video	360 × 288	MS-MPEG-4	Blockiness [21]	0.748
Mask	Image	768 × 512	JPEG2000	EHD [64]	0.792
		512 × 512	JPEG	MS-SSIM [64]	0.952
		512 × 512	JPEG2000	MSE [53]	0.913
		512 × 512	White noise	PSNR [53]	0.892
		480 × 720	JPEG,	W-SCM [53]	0.714
		512 × 512	JPEG,	PSNR [66]	0.688
		512 × 512	JPEG2000	PSNR [66]	0.724
		512 × 512	JPEG	SSIM [67]	0.731
		372 × 512	JPEG,	PSNR [69]	0.795
		372 × 512	JPEG2000	PSNR [69]	0.716
	Video	32 × 32(FFT)	VQEG	MSE [70]	0.821
		768 × 512	H.263	SSIM [73]	0.737
		768 × 512	H.264/AVC	EPSNR [21]	0.858
		486 × 720	MPEG-2	VQM [50]	0.923
		486 × 720	MPEG-2	PSNR Winkler Metric [73]	0.779
		32 × 32(FFT)	MPEG-2	Winkler Metric [71]	0.784
		32 × 32(FFT)	MPEG-2	Winkler Metric [74]	0.713

Resolution represents the image/video resolution. *Distortion type* represents their method used on different distortion types with respect to quality estimations. *Quality parameter* represents the evaluation parameters used by authors in their work.

Note: *Parameter value* represents performance values of *the RR technique* applying on particular distortion type. These values measure the amount of distorted introduced in the multimedia content. For example, if the distortion is due to the **JPEG** compression and someone want to measures the value by using **PSNR** with respect to a particular RR-based method. We convert and scale the parameter values in range from 0-1 used by authors in their RR-based method. If the parameter value is high, we can argue that the multimedia quality is less (i.e., highly degraded) and better will be the RR-based technique. The authors used these distortions types for RR technique to measure the metrics values (i.e., PSNR, SSIM)

extraction techniques as found in the National Telecommunications and Information Administration (NTIA) general Video Quality Model (VQM).

In [70], an objective VQA method is presented that uses the gain and loss of information of local harmonic strength for VQA. The harmonics information generated from edge-detected pictures is used to measure the quality degradations. First of all, edges information from the image is extracted, then an optional false edges removal operation is performed. The edges information is converted to blocks first and RR harmonics analysis is performed. This harmonics information is sent to the receiver side as the RR video quality measure. The receiver calculates the same harmonics information on the received video and then an objective comparison is used to check the quality degradation.

In [71, 72] the feature information is extracted using mask operations. In [21], the histogram difference and CBD is used for quality estimation which is basically a point-based operation. However, the structure information is extracted using mask-based operations as discussed in [50, 64, 66, 70].

A summary of mask-based RR perceptual quality assessment methods is presented in Table 1, with details. Mask-based approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance for better understanding.

4.1.3 Results and discussion

The information of pixel-based methods for both images and videos[7] multimedia contents can be obtained using point and mask-based operations. The main purpose of dividing the pixel-based methods to sub-classes is to reduce the search and time for those, who are interested in this area. The benefits of this classification are to select a particular domain for developing new techniques for RR-based methods and also explore only the relevant materials. Table 1 summarizes the reviews and literature work of pixel-based methods in detail with both the distortion type and related quality measurement parameters. For example, in order to develop a new technique for RR-based multimedia quality assessment and to use pixel-based approach, Table 1 provides a concrete and comprehensive literature review.

Table 1 also represents the comparison between point and mask-based approaches with respect to quality metrics and performance. Point-based techniques are mostly used for 2D-image quality estimations and the values used in the performance are based on the amount of how much the image is distorted by compression technique. The higher the values of a quality metrics (i.e., column no. 6), the higher is the multimedia quality degraded and better is the technique used for RR-based features. If we analyze in-depth Table 1 row 2, the values show JPEG2000 compression and PSNR for quality estimation with high value. Mask-based approaches are mostly used for video quality estimation and in few cases for the image. In Table 1, quality metrics and related literature (i.e., column-5) are displayed for selecting the relevant technique.

4.2 Frequency-based methods

Features in the form of frequency domain are important in all areas of multimedia data processing and analysis. The two most significant and commonly used feature extraction domains are wavelet and discrete cosine transformation (DCT) coefficient, which are widely used in the fields of multimedia analysis. *Wavelets* [75] are mathematical functions that cut up data into different frequency components, and study each component with a resolution matched to its scale. A wavelet series is a representation of a square-integrable (real- or complex-valued) function by a certain orthonormal series generated by a wavelet. The multimedia data represented in *DCT* [76] expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. In the field of RR I and VQA, researchers have developed different RR multimedia

---

[7] Here we are considering a video is the combinations of frames and their quality is in terms of individual-frame quality.

quality estimation approaches based on Wavelets and DCT features. Wavelet and DCT transformations are mathematical tools to utilize to abstract information from different kinds of multimedia including audio, images or videos. The frequency-based methods used DCT and wavelet coefficients to find the RR features. We have presented these approaches in detail in the following sections.

4.2.1 Discrete wavelet transform coefficient-based methods

Wavelet transform is a mathematical tool to extract information from the multimedia contents in the frequency domain. Wavelet transform is used to propose different RR multimedia contents quality assessment methods. Ranges of wavelets are usually needed to examine the data.

Wavelet-based watermarking schemes presented in [77–79] is the approximation of a parameterized Discrete Wavelet Transform (DWT). At the transmitter side, original image features are embedded in the original multimedia contents via a robust watermarking wavelet technique. In watermarking, the optimize parameter is DWT which is used to derive optimal wavelet transform for every image to solve the genetic algorithm optimization problem. After transmitting these embedded contents, at the receiver end, the entrenched features for the original multimedia contents are extracted and measured with the analogous features of the noisy multimedia contents. During transmission, the low-frequency feature contents of the multimedia and histogram features suffered distortion. These distortions are calculated by transmitting the low-level frequency features of original multimedia contents with their analogous histogram features at the receiver side by means of a strong watermark [80].

The image model which consists of natural image statistics is based on wavelet-domain with Kullback–Leibler [81] deviations. These features can be used as RR image and video quality measuring parameters for the quality estimation of image and video.

The multimedia contents quality proposed by [82] is a novel-based method. This method is based on wavelet-domain by using contrast sensitivity function (CSF), *multiscale geometric analysis*, and Weber's law of just noticeable difference to measure the RR features for image and video quality.

The statistic-based methods proposed by [83, 84] are based on a stochastic RR-IQA index to calculate the quality of the image by using state-of-the-art deep learning *Restricted Boltzmann Machine Similarity Measure*.

The RR perceptual quality assessment parameters proposed by [85–88] use RR perceptual quality metrics for color stereoscopic images. Their original and distorted RR model is presented in Fig. 5. Their approach is based on a disparity map of the original (reference) stereoscopic and distorted image, to measure the RR-based quality of multimedia content. In the first phase, the disparity map for reference and distorted image is measured by using color image disparity measurement by using eigenvalues and tensor structured properties of stereoscopic images. In the second phase, they measured multispectral wavelet decomposition to differentiate the different channels in the HVS. In the third phase, the CSF filtering is used to obtain the visual features information from both the reference and distorted images. Combining features from these stages, they estimated the RR features of distorted and reference stereoscopic multimedia quality and proposed RR IQA metrics for stereoscopic images.

Fig. 4 DCT-based RR image and video quality framework [76]
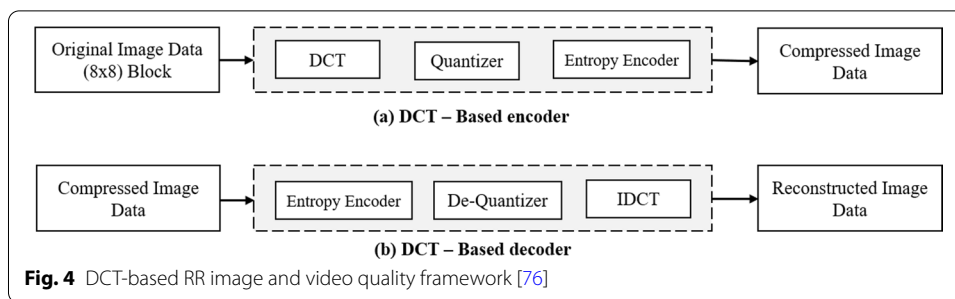
The work proposed by [89], is based on contrast sensitivity function-filtering as RR features for the reference image, and find the quality of the distorted image. By using the features of HVS, rational sensitivity threshing is set to extract the sensitivity coefficients of reference and distorted images and calculate the image quality parameter.

The low-level features of multimedia contents for quality measuring are proposed by [90, 91]. Their method is based on edge discrimination information in the form of image statistics. In the first phase, the binary edge map was measured from the wavelet-transform modulus and then multi-scale wavelet transformation of both reference and distorted images are measured to find the quality of multimedia content. The low-level features are used to differentiate between reference and distorted image features. In the second stage, the edge-pattern map is produced by using a gradient operator, which applied on the binary map called *edge-pattern map*. This edge-pattern map is further utilized to produce the histogram to verify the pattern of the edges of distorted and reference image to measure the quality of image-based on these patterns of edges.

Many researchers have used wavelet coefficient histograms as RR features in their work to compare with the distorted contents. The Kullback–Leibler, which used distance among the probability distributions of frequency-based wavelet coefficients of the references and distorted images and used it as a degree of multimedia contents distortion [92]. The second method, which is *Generalized Gaussian Model (GGM)* is used to summarize the peripheral dispersal of frequency features of wavelet coefficients for the reference image. In GGM, the assessment of multimedia contents quality need a small number of RR features. The spatio-temporal entropic differences [9] are linked well with human judgments of multimedia quality [93].

Table 2 shows the approaches proposed for discrete wavelet transform coefficient-based RR perceptual quality assessment methods. These approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance in detail.

### 4.2.2 Discrete cosine transform coefficient-based methods

Discrete cosine transformation (DCT) is used for the numerical solution of partial differential equations used in reference and distorted multimedia content analysis [76]. Spectral features of multimedia contents are done with the help of DCT [105]. The concept of DCT with respect to RR perceptual quality assessment is shown in Fig. 4.

DCT has been broadly accepted and engaged for the compression of multimedia contents, denoising, and deblocking. It is also used for RR features of image and video

Table 2 Summary of DWT and DCT-based RR perceptual quality assessment metrics values with respect to the distortion types (i.e., JPEG, JPEG2000). The authors used these distortions types for RR technique to measure the metrics values (i.e., PSNR, SSIM, etc.)

omain	Multimedia	Resolution	Distortion type	Quality parameter	Parameter value
DWT	Image	512 × 512	JPEG	SSIM[4]	0.752
		512 × 512	JPEG	MSE [77]	0.783
		512 × 512	JPEG	PSNR [83]	0.764
		512 × 720	JPEG	SSIM [84]	0.874
		512 × 720	JPEG2000	SSIM [85]	0.893
		512 × 720	JPEG	MSE[80]	0.694
		512 × 720	JPEG2000	MSE[80]	0.654
		512 × 512	JPEG2000	PSNR [87]	0.856
		512 × 512	White Gaussian Noise	MSSIM [88]	0.734
		512 × 512	JPEG	POROPOSE D [91]	0.844
	Video	486 × 620	MPEG	STRRED [92]	0.719
		486 × 620	H.264	SRRED[93]	0.724
DCT	Image	8 8 DCT	JPEG and MP3	PSNR [94]	0.871
		8 8 DCT	JPEG and MP3	WNSIM proposed[95]	0.738
		8 8 DCT	JPEG and JPEG2000	SSIM[96]	0.733
		512 × 512	JPEG, JPEG2000	PSNR [97]	0.876
		512 × 512	Blur	SSIM proposed [98]	0.704
		512 × 512	White Gaussian noise	PSNIR [99]	0.830
		768 × 512	JPEG, JPEG2000	VSNR [100]	0.626
		512 × 512	Blur	MS-SSIM proposed [94]	0.825
	Video	512 × 512	Additive white noise	[101]	0.789
		512 × 728	JPEG, MPEG-2 [102]	PSNR[103]	0.828
		512 × 512	**WGN**	**PSNR[104]**	**0.872**

Resolution represents the image/video resolution. *Distortion type* represents the method used by authors on distortion type with respect to quality estimations. *Quality parameter* represents the evaluation parameters used by authors in their work.
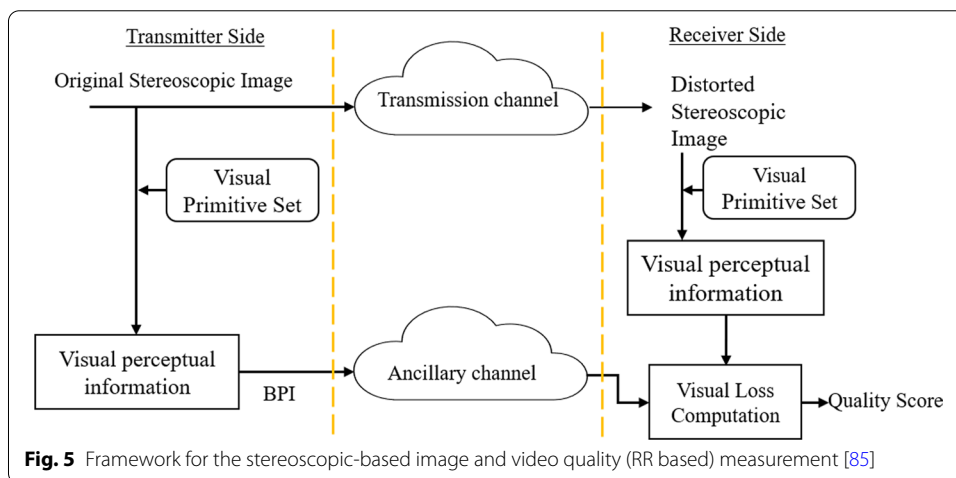
Fig. 5 Framework for the stereoscopic-based image and video quality (RR based) measurement [85]

quality measurement, analysis and even for guiding the image and video processing. We can abstract numerous HVS penetrating features based on the DCT coefficients. The approaches presented in [97, 106] are based on coefficient distributions of reduced DCT

subband, which can be precisely modeled by GGD for RR perceptual quality assessment. Those signals, which are sensitive to the addition of distortions can be used for the introduction of perceptual quality assessment methods [97]. These methods utilized the statistical properties of signals [98]. These signals' property will be affected by the introduction of distortion in multimedia content. As the image or video has been shown in the reduced DCT domain, the association between diverse frequency components should be measured and present. At the initial phase, the energy bending in diverse frequencies can marginally represent the alteration level. Secondly, HVS masking properties would be modeled by the energy distributions over different frequencies [99].

The approach presented in [101] used the magnitude and phase of 2D-DFT for the I and QA algorithms. The basic methodology is, to associate the magnitude and phase of reference and distorted images to calculate the quality parameters. To accommodate the fact that the human visual system [100] behaves differently for the different frequency elements. By enabling the RR features of the phase and frequency, the linear regression will combine the effects of changes in magnitude and phase. This technique is efficient and can be used to determine the required weights. This strategy used for RR perceptual quality assessment is phase-dependent due to the fact that phase carrying more information than magnitude.

The intra- and inter-subband statistical features in a simplified DCT domain are used for RR multimedia quality assessment [94, 104]. The approach in the intra-subband statistical features is the block-based DCT coefficients, which are simplified into a three-level tree. GGD function is used to capture the intra-subband characteristics. The main difference between the actual coefficient distribution and GGD is shown by City-Block Distance (CBD). In the second approach of inter-subband characteristics, the Mutual Information (MI) between adjacent reorganized DCT subband is used to show the corresponding relationships. The combination of the CBD intra-subband and MI of inter-subband depicts the proposed RR IQA in [104]. This method gives a very good result as compared to the existing RR methods which require a smaller number of RR features.

Table 2 shows the approaches proposed for DCT coefficient-based RR perceptual quality assessment methods. These approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance in detail.

4.2.3 Results and discussion
Literature review regarding frequency-based RR perceptual quality assessment methods are divided into wavelet and DCT-based features. The main purpose of dividing the frequency-based methods RR perceptual quality assessment into sub-classes is to reduce the search and time for those, who are interested in RR multimedia quality assessment and want to develop new parameters.

Wavelet transformation offers simultaneously in space and frequency a suitable framework for the limited representation of signals. They are mostly used to become the preferred form of demonstrations for many multimedia contents algorithms and used in the initial phases of biological visual-system [92].

The wavelet-based watermarking scheme is a parameterized DWT, which is an embedding process in the transmitter side to establish the wavelet-based watermarking within

the original images/frames. The features of the original image/frame will be secured and we will get DWT, which is an optimized parameter to solve the optimization of genetic algorithms. Another promising performance in wavelet-domain for multimedia content visual perception quality evaluation is grounded on the statistics property of the natural image. In wavelet-domain, the Natural Image Statistic Metric (NISM) is used to find the wavelet coefficient of original and degraded image/frame for multimedia quality assessment. These coefficients are represented in the GGD form, which essentially used for quality measurement [94]. The NISM method has been documented as the normal approach for RR multimedia contents quality, it fails to deliberate the statistical associations of wavelet coefficients in dissimilar subbands and the visual reaction characteristics of the mammalian cortical simple cells [107].

Wavelet conversion cannot openly abstract the image symmetrical and geometrical information, e.g., curves, lines, and wavelet coefficients are impenetrable to smooth multimedia contents edge contours. Consequently, there is a giant room for supplementary improvements in the efficiency of RR multimedia contents quality. To overcome these problems, the contourlet transformation [108] is a method for optimally representing image symmetrical and geometrical information. This method perceives, identifies, organizes and deploys data (e.g., lines, edges, and curves), which technically distance a high-dimensional space with significant features.

For decorrelation and energy compaction properties [109, 110], we have used DCT features because most of the research accomplishments in image and video codding have been engrossed on the DWT.

Table 2 summarizes the reviews and state-of-the-art approaches in the domain of frequency-based RR perceptual quality assessment methods in detail. Frequency feature approaches are presented in the image and video category. Table 2 also represents the results with respect to the distortion type of multimedia content and quality measurement parameters. For instance, to develop a new approach for RR-based multimedia quality assessment and to use frequency features as RR, Table 2 provides a tangible literature review. It will avoid the extra search for irrelevant approaches in this area.

4.3 Bitstream-based methods

Bitstream is the sequence of bits in the form of 1 s and 0 s, which can be transferred from one device (or location) to another. Bitstreams are used in communication, audio and networking applications. When we transferred the multimedia content (i.e., image or video) through a communication channel, there is a maximum probability that these media channels can add distortion. To find the amount of distortion and measure the quality of distorted multimedia content is a very important area of multimedia quality assessment research. We reported those approaches, which based on multimedia quality distortion with respect to communication media channels in the form of bitstream-based RR perceptual quality assessment methods.

Bitstream-based RR perceptual quality assessment methods make use of the bitstream data that are sent to the network channel. The encoders encode the original data and convert it to the bitstream which is then used in the bitstream-based VQA methods. This data is used for objective VQA, QoE and QoS for the end-users in the networked environment [22]. The stream of bits is parsed to get different features for the estimation
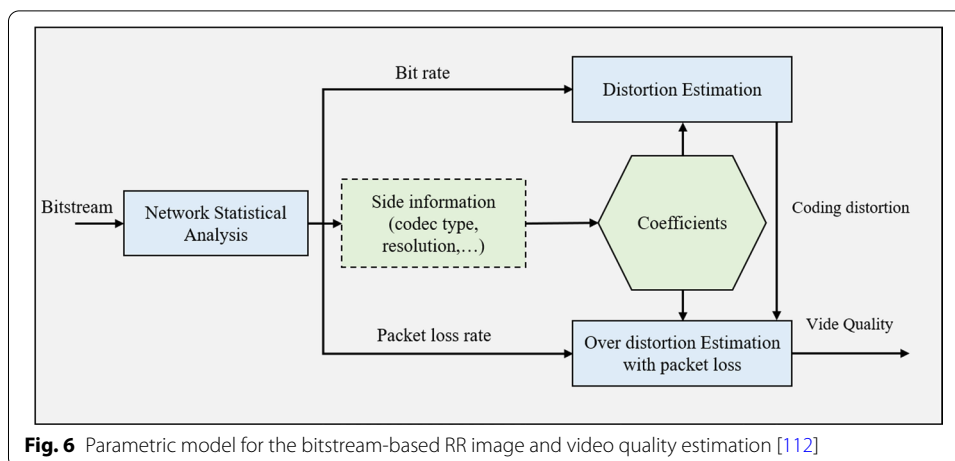
Fig. 6 Parametric model for the bitstream-based RR image and video quality estimation [112]

of quality. These methods are computationally less intensive as they do not need to decode the full encoded multimedia contents to estimate the quality [22].

Bitstream-based RR perceptual quality assessment methods are not universal as each network encoder uses different encoding standards. Hence the data of bitstream are in different formats [111]. In bitstream-based RR perceptual quality assessment methods, data loss occurs due to packet loss over the network. Video streaming services usually use User Datagram Protocol (UDP) and Real-time Transport Protocol (RTP) [111] as they do not cause unwanted delays. On the contrary, in Transmission Control Protocol (TCP) reliable data delivery is granted, but causes unwanted delays.

Communication channels can support either high or low bandwidth speed and transfer multimedia contents through these channels. Low bandwidth channel supports low data rate for transmission; which has high chances of data losses during transmission as compared to the high-bandwidth channel. High-bandwidth channel transmits huge amount of data simultaneously due to which the chances of data losses are minimum. Network applications that require high data transmission rate usually uses UDP and RTP transmission protocols as they do not cause unwanted delays. On the other hand, some applications require reliable data delivery which use TCP. TCP protocol causes unwanted delays either with low-bandwidth channels or high-bandwidth channels. Since the distortion of multimedia caused by each channel is different, we have classified the bitstream-based quality assessment methods into low- and high-bandwidth channels-based RR perceptual quality assessment methods [112]. A parametric model for bitstream-based RR perceptual quality assessment is presented in Fig. 6.

4.3.1 Low-bandwidth-based RR methods

The network performance is an important parameter for transmitting the data from one place (i.e., device) to another. The network performance affected by lower bandwidth also affects latency, jitter, packet loss, and throughput [113, 114]. These network performance parameters play a very important role in the quality of multimedia content. According to Hartley's law, the channel capacity of a physical communication link is proportional to its bandwidth [112, 115].

Fig. 7 Transmission channel used for the RR IQA deployment [116]
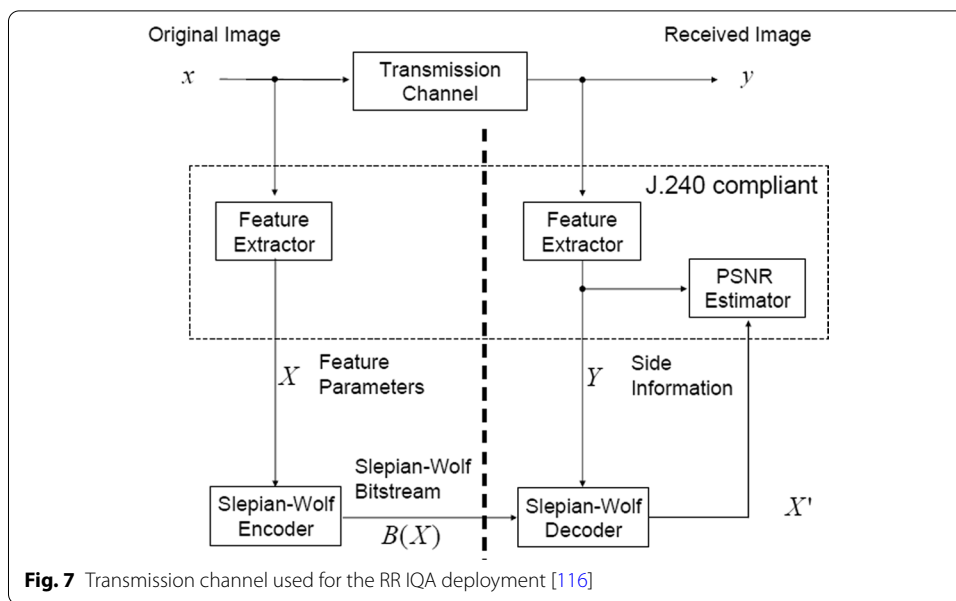
The category of low bandwidth channel bit-rate is in the range of 1 bit per second to 10 kilobits per second. We reported the work done in the field of RR perceptual quality assessment from 1G to 4.5G communication generation, because during the writing of this article the deployment and development of 5G are only on documents. First-generation (1G) communication bandwidth channels were mostly used for low-bitrate applications (i.e., instant messaging). Second-generation (2G) communication channels have provided image transmission capability and third-generation (3G) communication channels have offered enough bandwidth for transmission of digital images and videos. For HD videos and live video streaming, the bandwidth requirements are even higher, which is offered by fourth-generation (4G) and (4.5G) communication channels [113].

At the sender end in the communication channel, the reference multimedia content is first compressed and then transmitted to the channel. While at the receiver side the data are decompressed. During these three phases, three types of distortion are added to the multimedia contents: (1) compression, (2) transmission and (3) decompression. During transmission, a code is embedded with reference multimedia. At the receiver side, that code is extracted and compared with the original multimedia content and code to find the amount of distortion [116]. This approach is shown in Fig. 7.

The perceptual quality of image and video mostly relies on characteristic changes of both input reference multimedia contents and transmission channels. The bandwidth channel has different applications, i.e., small-size image transmission, video–telephonic, low-bitrate wireless imaging and digital broadcast television [117] applications.

ITU-T Recommendation G.1070 recommends a method for the videophone quality assessment, which based on video parameters and speech. Network performance organizations use this method to ensure the quality of services [118]. For low-bandwidth channels, the RR video quality evaluation system uses the reference features from the coarse video to find the quality of multimedia contents [119]. The system is tested for 18 subjectively rated data sets and it shows very good results for the low-bandwidth channel.

The system designed by [120] is based on RR IQA, which uses wireless imaging. The system used two types of RR features: (1) the *normalized hybrid image quality metric* and (2) perceptual relevance weighted Lp-norm for structural image information. The system first takes input image as a reference and find both image features for quality estimation and embed these features with the reference image. At the receiver end of the wireless system, the distorted image decodes for both features and find the image quality.

The quality parameters presented in [121–123] use together the National Telecommunications and Information Administration (NTIA) general Image and Video Quality Model (IVQM) to map 19 subjective data sets into a F, T (i.e., F=false,T=true) subjective quality scale [121]. The resulting subjective data were used to find the most suitable linear combination of the 9 video quality parameters in the 10-kbps IVQM.

The quality of the video in IPTVs is measured in ITU-T J.240 [124] using PNSR. The RR method for subjective VQA in [45] measures video quality, but cannot handle the errors introduced by the transmission procedure. VQA methods that deal with the video degradation caused by transmission and video compression are under study at Video Quality Experts Group (VQEG) [116].

The method proposed in [125] uses activity difference between original and transmitted video to estimate the quality. The activity difference of the original video is calculated at the sender side and is transmitted to the receiver side along with the original video. The activity difference of the received video is calculated at the receiver and is compared with the activity difference of the original video for quality estimation. The method also uses a weighting for the activity difference values. For example, in a video frame, the region of interest is a human being. The pixel values greater than 175 approximate to the human skin color and is multiplied by a constant weight. In the same way, high frequencies are also given predefined weights as HVS is less sensitive to high frequencies; so the weights reduce the effect of high frequencies [125]. The approach in [125] uses temporal sub-sampling and partial bit information transmission, i.e., lower 6 bits are transmitted as they contain more information. After channel degradations, original and degraded videos are highly correlated, due to this reason only lower bits are sent as RR for VQA. The method is tested against subjectively tested video quality [126] results.

The method proposed in [80] uses a low-frequency coefficient and low histogram information of the original image as a RR feature for estimating the channel-induced errors. This information is embedded in the original image as a watermark. In practice, there is no ancillary channel available all the time to send RR features independent of the original image. Therefore, low-frequency coefficients of the image are calculated using 2D wavelet transform and are embedded as a watermark in the original image. The original image with an embedded watermark is sent over the network. On the receiver side, a 2D wavelet transform is applied to the distorted image and the watermark is extracted.

Table 3 shows the approaches proposed for low-bandwidth channel-based RR perceptual quality assessment methods. These approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance in detail.

Fig. 8 High-bandwidth channel video quality monitoring-based on an RR model [127]
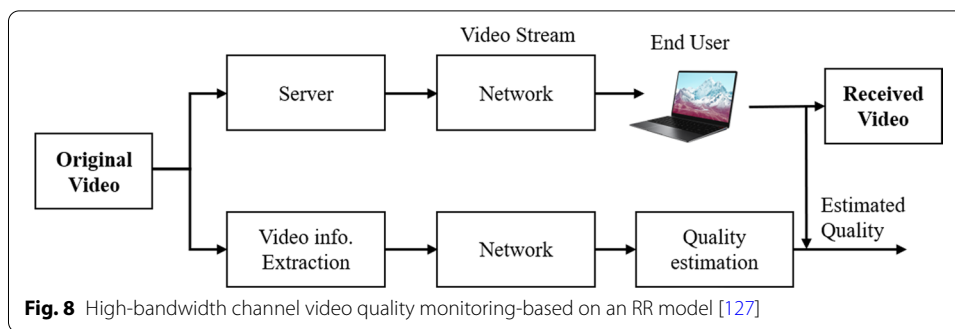
4.3.2 High-bandwidth-based RR methods

The range of high-bandwidth channel starts from 10 kbps. Different applications of high bandwidth require different bandwidth ranges. (i) voice over IP (VoIP) requires 56.5 kbps to transmit sound clearly and smooth; (ii) standard definition video (481p) work at 2 megabit per second (Mbps); (iii) HD video (740p) requires more than 5 Mbps, and (iv) HD XenDesktop (HDX) (1080p) requires more than 8 Mbps [127]. When videos are transmitted over the channel, some unwanted features(noise) are added with the received video. The objective assessment for measuring the quality of the received video is a subject of great importance. PSNR was previously used as an objective parameter for VQA, but their results have poor correlation with HVS response to visual quality [128]. Another parameter used in [129] exploits video structural similarity index (VSSI) [130] for VQA. The results of VSSI have a good correlation with the subjective measure of MOS [80]. High-bandwidth channel video quality system of RR is shown in Fig. 8.

The RR methods that use low data for RR information are non-linear quantization [66] and distributed source coding [131]. RR methods based on high bandwidth can be either designed autonomously with respect to already existing FR methods [47] or as an approximation of some FR metrics as in [66].

The method proposed in [129] uses a feature metrics to estimate visual quality. The feature metrics of the original video is extracted at the sender's side and sent to the receiver over a noiseless channel. The original video is sent to the receiver after encoding; the video is decoded at the receiver end and the feature metrics are estimated by NR means. Then the structural similarity of the estimated feature metrics and received feature metrics (received by noiseless channel) is measured to estimate the visual quality.

The approaches presented in [132, 133] used comparison with respect to structural degradation index in multimedia contents. These approaches are tested for different compression ratios and network situations. Moreover, it uses less information to be used as a reference for measuring quality distortions.

Table 3 shows the approaches proposed for high-bandwidth channel-based RR perceptual quality assessment methods. These approaches are presented in the image and video category. Each approach is described with the multimedia content resolution, processing, quality parameter (i.e., metrics) and parameter performance in detail.

4.3.3 Results and discussion

The state-of-the-art approaches in the field of RR perceptual quality assessment methods with respect to bitstream-based techniques for multimedia contents can be divided into low- and high-bandwidth channels. The main purpose of dividing bitstream methods into sub-classes is to reduce the search and time for those who are interested in RR perceptual quality assessment developing and analysis.

In signal processing and communication, bitstream-based perceptual quality assessment methods are used for objective VQA, QoE and QoS [22, 111]. The quality of these methods are computationally less intensive [22, 134]. Moreover, these methods are not universal, because the data of bitstream are in different formats. Better distortion estimation is achieved in RR techniques as compared to NR techniques due to the availability of original bitstream at the receiver side.

Another significance of bitstream perceptual quality assessment methods is their computational simplicity, which plays an important role in the online quality monitoring systems. Bitstream perceptual quality assessment methods have distinct gain over the pixel and frequency-based methods due to the availability of access to the core bit-rate, frame per second, quality of service and different types of features. These impact and degrade the quality of the network.

For the computational load, the parametric packet-layer models are very useful by utilizing the in-service nonintrusive QoE measurement [135]. Due to this reason, they do not utilize the capacity to look at the payload information. The measurement of QoE of individual users is not possible in real-world scenarios in which transmitter encodes the RR features with multimedia contents. To solve this problem, the coded bitstream information is used to utilize the characteristics of source feature. This scenario can be used in DCT coefficients in MPEG coding, which tell us about the spatial complexity of the multimedia contents region of interest [136].

Packet-layer models do not need decryption at the receiver side, due to which its performance is better and which makes it popular to use in network applications. Bitstream-layer models are more efficient than packet-layer models on the basis of performance and complexity. Bitstream-layer models due to its dynamic and flexible nature are used for the desired level of achieving accuracy. The property of flexibility makes bitstream-based methods superior than pixel-based or frequency-based methods.

Table 3 summarizes the reviews and literature work of bitstream-based methods in detail with both the distortion type and related quality measurement parameters. For example, to develop a new technique for RR-based multimedia quality assessment and to use a bitstream-based approach, Table 3 provides a concrete and comprehensive literature review. Furthermore, for finding the in-depth technique used for bitstream-based methods, Table 3 sub-classes (column-1) provide that information without wasting time in exploring other techniques [134].

Table 3 also presents the comparison between high- and low-bandwidth-based RR approaches with respect to quality metrics and performance. Low-bandwidth-based techniques are mostly used for video quality estimations and the values used in the performance are based on the amount of how much the multimedia content is distorted by compression technique. The higher the values of performance quality metrics (column-6), the higher is the quality of multimedia content and the better the technique for

Fig. 9 High-bandwidth channel used for the video quality assessment using RR technique [131]
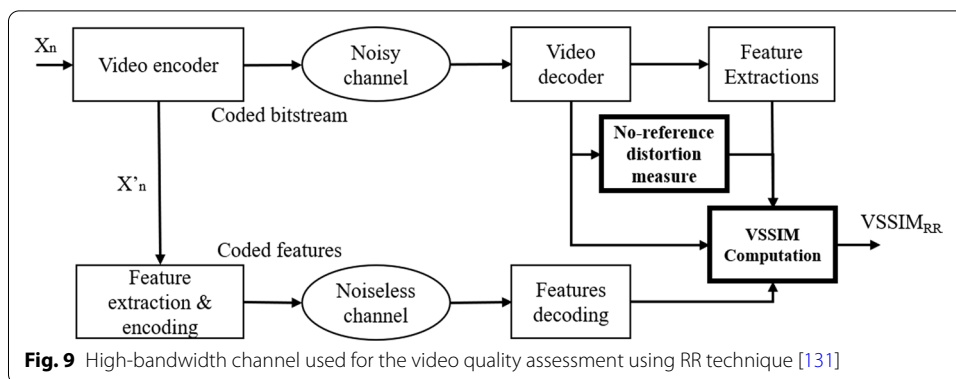
Table 3 Review of the bitstream-based RR I and VQA metrics performance values with respect to the distortion types (JPEG, JPEG2000, etc.) used by authors in their RR technique to measure the quality metrics values (PSNR, SSIM, etc.)

Domain	Multimedia	Resolution	Distortion type	Quality parameter	Parameter value
Low bandwidth	Image	16 × 16 pixel block	JPEG 2000	PSNR [122]	0.815
		512 × 512	JPEG 2000	MSSIM [123]	0.873
		512 × 512	JPEG	PSNR [128]	0.783
		16 × 16 pixel block	JPEG	MSSIM [121]	0.877
	Video	486 line × 720 pixel	VQEG	VQM [119]	0.724
		472 line × 480 pixel	FRTV	VQM [71, 120]	0.745
		VQEG FR-TV	MPEG-1	RMSE MOS [116]	0.728
		VQEG FR-TV	MPEG-2	PSNR [124]	0.748
		Phase-1 525-line	MPEG-2	PSNR [125]	0.865
		720 × 480	VQEG	PSNR [137]	0.809
High bandwidth	Image	512 × 512	JPEG, JPEG2000	MSSIM [131]	0.847
		512 × 512	JPEG	SSIM [136]	0.789
		704 × 576	JPEG2000	PSNR [138]	0.782
		720 × 576	JPEG	PSNR [132]	0.842
		720 × 576	JPEG2000	PSNR [134]	0.857
		512 × 512	JPEG2000	MSSIM [9]	0.828
	Video	486 × 620	H.264[139]	VSSI [129]	0.852
		486 × 620	H.264	AVC [9]	0.778
		720 × 576	H.264/AVC	SSIM [133]	0.887
		720 × 576	MPEG4	MSSIM [135]	0.842
		512 × 512	VQEG	PSNR [130]	0.8042
		512 × 728	JPEG, MPEG-2	PSNR [137]	0.828
		512 × 512	H.264/AVC	PSNR [139]	0.812

In the table, *Resolution* is used to represent the multimedia content (images and videos in 2D and 3D) resolution. *Distortion type* means to represent the techniques used by authors on distortions type in their work for multimedia content quality estimations. *Quality parameter* column represents the evaluation parameters used by authors in their respective methods

that particular RR features. If we look closely at the table, the value in bold in the seventh row shows JPEG compression and MSSIM for quality estimation with the highest value, indicating the best RR technique in low-bandwidth channel-based approaches.

High-bandwidth channel-based techniques are used for both image and video quality estimations and the values used in the performance are based on the amount of how

Fig. 10 Framework for the 3D RR-IQA [140]
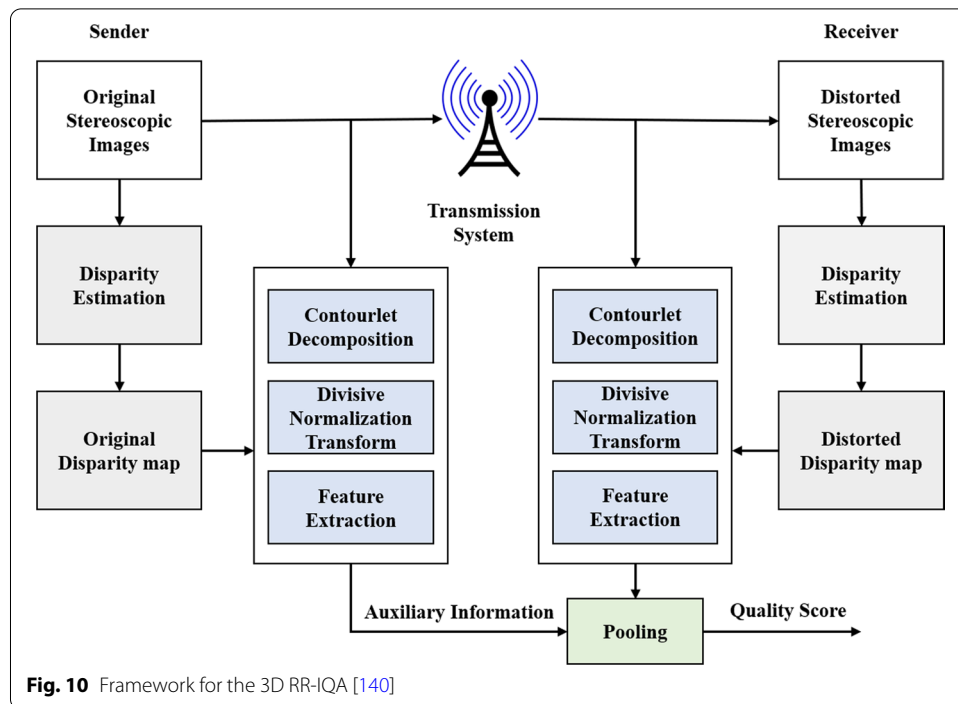
much the multimedia content is distorted by compression and transmission phase. If we look closely at the table, the value in bold in the sixth row shows H.264 compression and SSIM for quality estimation with the highest value indicating the best RR technique in high-bandwidth channel-based approaches. The details of high-bandwidth channel used for the video quality assessment using RR are shown in Fig. 9.

4.4 Three-dimensional (3D)-based method

The designing of reliable three-dimensional (3D) RR-IQA metrics is the future direction of full, reduced and no-reference image quality estimation and challenging task due to the exact calculation of image features in the 3D domain. A novel 3D RR-IQA-based approach is proposed in [140], which is based on the *Gaussian scale mixtures* model to normalized the coefficient in the contourlet subband [110, 141, 142] of image luminance and map disparity of 3D images. In [140], the feature similarity index with fitted Gaussian distribution is used to determine the feature similarity of RR features of a reference image for quality estimation. The 3D features are embedded in the reference image and pass through distorted media, as shown in Fig. 10. At the receiver end, the 3D RR-IQA on distorted image is decoded and the 3D features for quality estimation with embedded 3D features are found.

After 3D RR-IQA, the second direction for future research in the area of full, reduced and no-reference quality measurement will be 3D RR video quality assessment (3D RR-VQA). In order to provide a better quality of the 3D video to the customer will be the demanding issue of future multimedia applications. The quality measurement of 3D video will be a challenging task for researchers in the future for (1) multimedia applications, (2) quality of online services measurement, and (3) online 3D video streaming quality measurement. The work proposed for 3D RR-VQA by [143–145] is based on

color and depth information of the maps and similarly for information from the reference 3D video to measure the quality of the 3D video.

The method proposed in [132] presents a way of measuring 3D video transmission and compression degradations using the RR technique. As the original video is not available at the receiver side to measure the quality degradation, the RR feature of the original video is calculated and sent to the receiver for VQA. The RR feature used in this approach is 3D video color and depth map. The information provided by depth map coincides with the edges and contours information.

The comparison of the original and degraded image depth map and edges information estimates the quality degradations. The last part of Table 1 described the approaches related to 3D image and video-based methods RR perceptual quality assessment with their multimedia contents resolution, processing, proposed quality metrics and the performance of quality metrics.

5 Conclusions

The RR methods for perceptual quality assessment of images and videos are used in many practical applications of multimedia. The demand of these parameters for quality assessment will increase in the future with the deployment of the fifth-generation network (i.e., 5G). In this paper, we present a review of the RR-based multimedia quality assessment methods (approaches) and classify these methods into sub-classes (subdomain) on the basis of multimedia content processing. We also present the databases used for the development and evaluations of the RR parameters. We divided the RR-based methods into *pixel*, *frequency*, *bitstream*, and *3D multimedia*-based methods, which are unique and meaningful with respect to their content-based interpretation. The *Pixel*-based methods are the mainstream, which uses the multimedia contents pixel values as an input to the quality assessment algorithms. The *frequency*-based RR methods use the frequency-transformed features for quality estimation in reference and distorted multimedia contents. The *bitstream*-based RR methods are developed to measure the amount of distortion introduced over communication channel of the multimedia. At the end, we also present RR-based assessment methods developed for the *3D multimedia* quality measurement. RR-based methods provided a great potential over FR and NR methods [146–148]. These methods can be applied to practical scenarios, while FR and NR-based approaches are not suitable.

In most of the scenarios, RR methods are preferred, while FR methods require large amount of data to be processed for the quality assessment.

Since NR methods are blind methods which use only extracted information of the original signal to assess the quality; the quality assessed by NR methods might not be reliable. Finally, we conclude that this article would be useful to gain in-depth knowledge and jump-start development in the area of RR-based perceptual quality assessment methods.

Abbreviations
FR: Full reference; RR: Reduced reference; NR: No-reference; QoE: Quality of experience; QoS: Quality of service; MOS: Mean opinion score; QA: Quality assessment; HVS: Human visual system; NVC: Natural visual characteristics; LHS: Local harmonic strength; DCT: Discrete cosine transform; IQA: Image quality assessment; VQA: Video quality assessment; IVC: Image and video communication; DFT: Discrete Fourier transform; MSE: Mean square error; PSNR: Peak signal-to-noise ratio; OS: Orientation selectivity; PC: Phase congruency; EVD: Energy variation descriptor; GGD: Generalized Gaussian

Dost *et al. EURASIP Journal on Image and Video Processing*   (2022) 2022:1

Page 27 of 31

density; CBD: City-block distance; SSIM: Structural similarity; SVD: Singular value decomposition; MS-SSIM: Multi-scale structural similarity; FJLT: Fast Johnson–Lindenstrauss; NTIA: National Telecommunications and Information Administration; VQM: Video quality model; 3D: Three dimensional; RR-VQA: Reduced reference video quality assessment; RR-IQA: Reduced reference image quality assessment; DWT: Discrete wavelet transform; CSF: Contrast sensitivity function; GGM: Generalized Gaussian model; MI: Mutual information; UDP: User Datagram Protocol; RTP: Real-time Transport Protocol; TCP: Transmission Control Protocol; IVQM: Image and video quality model; VQEG: Video Quality Expert Group; VoIP: Voice over IP; mbps: Megabit per second; HDX: HD xenDesktop; VSSI: Video Structural Similarity Index.

Authors' contributions
SD prepared the research papers database in field of reduced reference image and video quality, rationalized the publications to be reviewed in this paper. He also planned the architecture of paper with the assistance of co-authors and wrote the main draft of the paper with FS and MS. FS and MS contributed in the writing first draft of the paper and involved in discussions to improve the quality. MG was involved in the topic refining and discussion of the paper, he also contributed in the critical, formatting, and review of paper in many stages. MS was involved in the topic selection and initial discussion of the paper, he also contributed in the review of paper. BL was involved in the initial discussion, review and quality checking of paper. All authors read and approved the final manuscript.

Availability of data and materials
Not applicable.

Declarations

Competing interests
Not applicable.

Author details
[1]TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany. [2]Department of Computer Science, National University of Computer and Emerging Sciences, Faisalabad, Pakistan. [3]Department of Electrical Engineering, National University of Computer and Emerging Sciences, Faisalabad, Pakistan. [4]Blekinge Institute of Technology, Blekinge, Sweden. [5]The Blekinge Institute of Technology, Karlskrona, Sweden.

References
1. M.G. Martini, B. Villarini, F. Fiorucci, A reduced-reference perceptual image and video quality metric based on edge preservation. EURASIP J. Adv. Signal Process. **2012**(1), 66 (2012)
2. Z. Wang, A.C. Bovik, Reduced-and no-reference image quality assessment. IEEE Signal Process. Mag. **28**(6), 29–40 (2011)
3. M. Carnec, P. Le Callet, D. Barba, Full reference and reduced reference metrics for image quality assessment. In: Proceedings of the Seventh International Symposium on Signal Processing and Its Applications, IEEE, vol. 1, pp. 477–480 (2003)
4. Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
5. C. Li, A.C. Bovik, Content-partitioned structural similarity index for image quality assessment. Signal Process. **25**(7), 517–526 (2010)
6. L. Zhang, L. Zhang, X. Mou, D. Zhang, Fsim: A feature similarity index for image quality assessment. IEEE Trans. Image Process. **20**(8), 2378–2386 (2011)
7. W. Xue, L. Zhang, X. Mou, A.C. Bovik, Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. IEEE Trans. Image Process. **23**(2), 684–695 (2013)
8. H. Liu, C. Li, D. Zhang, Y. Zhou, S. Du, Enhanced image no-reference quality assessment based on colour space distribution. IET Image Process. **14**(5), 807–817 (2020)
9. R. Soundararajan, A.C. Bovik, Rred indices: Reduced reference entropic differencing for image quality assessment. IEEE Trans. Image Process. **21**(2), 517–526 (2011)
10. C.G. Bampis, P. Gupta, R. Soundararajan, A.C. Bovik, Speed-qa: Spatial efficient entropic differencing for image and video quality. IEEE Signal Process. Lett. **24**(9), 1333–1337 (2017)
11. L.S. Chow, H. Rajagopal, Modified-brisque as no reference image quality assessment for structural mr images. Magn. Reson. Imaging **43**, 74–87 (2017)
12. Y. Zhang, A.K. Moorthy, D.M. Chandler, A.C. Bovik, C-diivine: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes. Signal Process. **29**(7), 725–747 (2014)

13. T. Goodall, A.C. Bovik, No-reference task performance prediction on distorted lwir images. In: 2014 Southwest Symposium on Image Analysis and Interpretation, IEEE, pp. 89–92 (2014)

14. H.Z. Nafchi, M. Cheriet, Efficient no-reference quality assessment and classification model for contrast distorted images. IEEE Trans. Broadcast. **64**(2), 518–523 (2018)

15. P. Ye, J. Kumar, L. Kang, D. Doermann, Unsupervised feature learning framework for no-reference image quality assessment. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1098–1105 (2012)

16. X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, C.W. Chen, Blind quality assessment based on pseudo-reference image. IEEE Trans. Multimedia **20**(8), 2049–2062 (2017)

17. X. Min, G. Zhai, K. Gu, Y. Liu, X. Yang, Blind image quality estimation via distortion aggravation. IEEE Trans. Broadcast. **64**(2), 508–517 (2018)

18. K. Gu, G. Zhai, X. Yang, W. Zhang, Using free energy principle for blind image quality assessment. IEEE Trans. Multimedia **17**(1), 50–63 (2014)

19. K. Gu, G. Zhai, X. Yang, W. Zhang, L. Liang, No-reference image quality assessment metric by combining free energy theory and structural degradation model. In: 2013 IEEE International Conference on Multimedia and Expo (ICME), IEEE pp. 1–6 (2013)

20. B. Ciubotaru, G.-M. Muntean, G. Ghinea, Objective assessment of region of interest-aware adaptive multimedia streaming quality. IEEE Trans. Broadcast. **55**(2), 202–212 (2009)

21. S. Winkler, A. Sharma, D. McNally, Perceptual video quality and blockiness metrics for multimedia streaming applications. In: Proceedings of the International Symposium on Wireless Personal Multimedia Communications, pp. 547–552 (2001)

22. M. Shahid, A. Rossholm, B. Lövström, H.-J. Zepernick, No-reference image and video quality assessment: a classification and review of recent approaches. EURASIP J. Image Video Process. **2014**(1), 40 (2014)

23. T. Zhu, L. Karam, A no-reference objective image quality metric based on perceptually weighted local noise. EURASIP J. Image Video Process. **2014**(1), 5 (2014)

24. R. Soundararajan, A.C. Bovik, Survey of information theory in visual quality assessment. Signal Image Video Process. **7**(3), 391–401 (2013)

25. Z. Wang, A.C. Bovik, Modern image quality assessment. Synth. Lect. Image Video Multimedia Process. **2**(1), 1–156 (2006)

26. S. Chikkerur, V. Sundaram, M. Reisslein, L.J. Karam, Objective video quality assessment methods: A classification, review, and performance comparison. IEEE Trans. Broadcast. **57**(2), 165–182 (2011)

27. S. Wang, X. Zhang, S. Ma, W. Gao, Reduced reference image quality assessment using entropy of primitives. In: Picture Coding Symposium (PCS), IEEE, 2013, pp. 193–196 (2013)

28. Y. Niu, Y. Zhong, W. Guo, Y. Shi, P. Chen, 2d and 3d image quality assessment: A survey of metrics and challenges. IEEE Access **7**, 782–801 (2018)

29. U. Engelke, H.-J. Zepernick, Perceptual-based quality metrics for image and video services: A survey. In: 3rd EuroNGI Conference on Next Generation Internet Networks, IEEE, pp. 190–197 (2007)

30. W. Lin, C.-C.J. Kuo, Perceptual visual quality metrics: A survey. J. Visual Commun. Image Represent. **22**(4), 297–312 (2011)

31. S. Winkler, P. Mohandas, The evolution of video quality measurement: From psnr to hybrid metrics. IEEE Trans. Broadcast. **54**(3), 660–668 (2008)

32. A.M. Eskicioglu, P.S. Fisher, Image quality measures and their performance. IEEE Trans. Commun. **43**(12), 2959–2965 (1995)

33. H.R. Sheikh, A.C. Bovik, G. De Veciana, An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans. Image Process. **14**(12), 2117–2128 (2005)

34. A. Shnayderman, A. Gusev, A.M. Eskicioglu, An svd-based grayscale image quality measure for local and global assessment. IEEE Trans. Image Process. **15**(2), 422–429 (2006)

35. D.M. Chandler, S.S. Hemami, Vsnr: A wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans. Image Process. **16**(9), 2284–2298 (2007)

36. H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Process. **15**(11), 3440–3451 (2006)

37. S. Xu, S. Jiang, W. Min, No-reference/blind image quality assessment: a survey. IETE Tech. Rev. **34**(3), 223–245 (2017)

38. Y. Niu, L. Lin, Y. Chen, L. Ke, Machine learning-based framework for saliency detection in distorted images. Multimedia Tools Appl. **76**(24), 26329–26353 (2017)

39. D.M. Chandler: Seven challenges in image quality assessment: Past, present, and future research. ISRN Signal Process. **2013** (2013)

40. Z. Guangtao, M. Xiongkuo, Perceptual image quality assessment: A survey. SCIENCE CHINA Information Sciences

41. H. Sheikh, Live image quality assessment database release 2. http://live.ece.utexas.edu/research/quality (2005)

42. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, Tid 2008-a database for evaluation of full-reference visual quality assessment metrics. Adv. Modern Radioelectron. **10**(4), 30–45 (2009)

43. N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti et al., Image database tid2013: Peculiarities, results and perspectives. Signal Process. **30**, 57–77 (2015)

44. S. Corchs, F. Gasparini, R. Schettini, No reference image quality classification for jpeg-distorted images. Digit. Signal Process. **30**, 86–100 (2014)

45. P. Le Callet, F. Autrusseau, Subjective quality assessment irccyn/ivc database (2005)

46. K. Seshadrinathan, R. Soundararajan, A.C. Bovik, L.K. Cormack, A subjective study to evaluate video quality assessment algorithms. Hum. Vision Electron. Imaging **7527**, 75270 (2010)

47. Z. Wang, L. Lu, A.C. Bovik, Video quality assessment based on structural distortion measurement. Signal Process. **19**(2), 121–132 (2004)

48. Y. Horita, Image quality evaluation database. ftp://guest@ mict. eng. u-toyama. ac. jp/ (2000)

49. M. Carnec, P. Le Callet, D. Barba, Visual features for image quality assessment with reduced reference. In: IEEE International Conference on Image Processing, 2005. ICIP 2005, IEEE. vol. 1, p. 421 (2005)

50. D.-O. Kim, R.-H. Park, D.-G. Sim, Reduced-Reference Image Quality Assessment Based on the Similarity of Edge Projections

51. L. Goldmann, F. De Simone, T. Ebrahimi, Impact of acquisition distortion on the quality of stereoscopic images. In: Proceedings of the International Workshop on Video Processing and Quality Metrics for Consumer Electronics (2010)

52. S. Winkler, Analysis of public image and video databases for quality assessment. IEEE J. Select. Top. Signal Process. **6**(6), 616–625 (2012)

53. W. Xue, X. Mou, Reduced reference image quality assessment based on weibull statistics. In: Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on IEEE, pp. 1–6 (2010)

54. J. Wu, W. Lin, G. Shi, A. Liu, Reduced-reference image quality assessment with visual information fidelity. IEEE Trans. Multimedia **15**(7), 1700–1705 (2013)

55. J. Wu, W. Lin, G. Shi, L. Li, Y. Fang, Orientation selectivity based visual pattern for reduced-reference image quality assessment. Inf. Sci. **351**, 18–29 (2016)

56. D. Liu, Y. Xu, Y. Quan, P. Le Callet, Reduced reference image quality assessment using regularity of phase congruency. Signal Process. **29**(8), 844–855 (2014)

57. Y. Xu, D. Liu, Y. Quan, P. Le Callet, Fractal analysis for reduced reference image quality assessment. IEEE Trans. Image Process. **24**(7), 2098–2109 (2015)

58. K. Okarma, P. Lech, A statistical reduced-reference approach to digital image quality assessment. In: International Conference on Computer Vision and Graphics, Springer pp. 43–54 (2008)

59. M. Carnec, P. Le Callet, D. Barba, An image quality assessment method based on perception of structural information. In: Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on IEEE, vol. 3, p. 185 (2003)

60. K. Gu, G. Zhai, X. Yang, W. Zhang, A new reduced-reference image quality assessment using structural degradation model. In: Circuits and Systems (ISCAS), 2013 IEEE International Symposium on IEEE, pp. 1095–1098 (2013)

61. Q. Li, Z. Wang, General-purpose reduced-reference image quality assessment based on perceptually and statistically motivated image representation. In: Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on IEEE, pp. 1192–1195 (2008)

62. A. Rehman, Z. Wang, Reduced-reference image quality assessment by structural similarity estimation. IEEE Trans. Image Process. **21**(8), 3378–3389 (2012)

63. P. Paudyal, F. Battisti, M. Carli, Reduced reference quality assessment of light field images. IEEE Trans. Broadcast. **65**(1), 152–165 (2019)

64. E. Kalatehjari, F. Yaghmaee, Using structural information for reduced reference image quality assessment. In: Computer and Knowledge Engineering (ICCKE), 2014 4th International eConference on IEEE, pp. 537–541 (2014)

65. L. Liu, T. Wang, H. Huang, A.C. Bovik, Video quality assessment using space-time slice mappings. Signal Process. **82**, 115749 (2020)

66. X. Lv, Z.J. Wang, Reduced-reference image quality assessment based on perceptual image hashing. In: Image Processing (ICIP), 2009 16th IEEE International Conference on IEEE, pp. 4361–4364 (2009)

67. J.A. Redi, P. Gastaldo, I. Heynderickx, R. Zunino, Color distribution information for the reduced-reference assessment of perceived image quality. IEEE Trans. Circ. Syst. Video Technol. **20**(12), 1757–1769 (2010)

68. C. Charrier, G. Lebrun, O. Lezoray, A color image quality assessment using a reduced-reference image machine learning expert. In: Image Quality and System Performance V, International Society for Optics and Photonics vol. 6808, p. 68080 (2008)

69. M. Yu, H. Liu, Y. Guo, D. Zhao, A method for reduced-reference color image quality assessment. In: CISP'09. 2nd International Congress on Image and Signal Processing, pp. 1–5 (2009)

70. I.P. Gunawan, M. Ghanbari, Reduced-reference video quality assessment using discriminative local harmonic strength with motion consideration. IEEE Trans. Circ. Syst. Video Technol. **18**(1), 71–83 (2008)

71. Q. Li, Z. Wang, Reduced-reference image quality assessment using divisive normalization-based image representation. IEEE J. Select. Top. Signal Process. **3**(2), 202–211 (2009)

72. S. Dost, S. Anwer, F. Saud, M. Shabbir, Outliers classification for mining evolutionary community using support vector machine and logistic regression on azure ml. In: International Conference on Communication, Computing and Digital Systems (C-CODE), IEEE pp. 216–221 (2017)

73. Z. Wang, A.C. Bovik, *Modern Image Quality Assessment (Synthesis Lectures on Image, Video, and Multimedia Processing)* (Morgan Claypool, San Rafael, 2006)

74. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2004, IEEE, vol. 2, pp. 1398–1402 (2003)

75. Y. Meyer, Wavelets-algorithms and applications. Wavelets-Algorithms and applications Society for Industrial and Applied Mathematics Translation., 142 p. (1993)

76. N. Ahmed, T. Natarajan, K.R. Rao, Discrete cosine transform. IEEE Trans. Comput. **100**(1), 90–93 (1974)

77. A.N. Avanaki, S. Sodagari, A. Diyanat, Reduced reference image quality assessment metric using optimized parameterized wavelet watermarking. In: Signal Processing, 2008. ICSP 2008. 9th International Conference On, IEEE, pp. 868–871 (2008)

78. W. Lu, X. Gao, D. Tao, X. Li, A wavelet-based image quality assessment method. Int. J. Wavelets Multiresolution Inf. Process. **6**(04), 541–551 (2008)

79. S. Altous, M.K. Samee, J. Götze, Reduced reference image quality assessment for jpeg distortion. In: ELMAR, 2011 Proceedings, IEEE pp. 97–100 (2011)

80. M.H. Kayvanrad, S. Sodagari, A.N. Avanaki, H. Ahmadi-Noubari, Reduced reference watermark-based image transmission quality metric. In: Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium On, IEEE pp. 526–531 (2008)

81. J.R. Hershey, P.A. Olsen, Approximating the kullback leibler divergence between gaussian mixture models. In: 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, IEEE vol. 4, p. 317 (2007)

82. X. Gao, W. Lu, D. Tao, X. Li, Image quality assessment based on multiscale geometric analysis. IEEE Transactions on Image Processing **18**(7), 1409–1423 (2009)

83. D.C. Mocanu, G. Exarchakos, H.B. Ammar, A. Liotta, Reduced reference image quality assessment via boltzmann machines. In: Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium On, IEEE pp. 1278–1281 (2015)

84. S. Bosse, D. Maniry, T. Wiegand, W. Samek, A deep neural network for image quality assessment. In: Image Processing (ICIP), 2016 IEEE International Conference On, IEEE, pp. 3773–3777 (2016)

85. A. Maalouf, M.-C. Larabi, Cyclop: A stereo color image quality assessment metric. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference On, IEEE, pp. 1161–1164 (2011)

86. P. Campisi, P. Le Callet, E. Marini, Stereoscopic images quality assessment. In: Signal Processing Conference, 2007 15th European, IEEE pp. 2110–2114 (2007)

87. F. Qi, D. Zhao, W. Gao, Reduced reference stereoscopic image quality assessment based on binocular perceptual information. IEEE Trans. Multimedia **17**(12), 2338–2344 (2015)

88. W. Zhou, G. Jiang, M. Yu, F. Shao, Z. Peng, Reduced-reference stereoscopic image quality assessment based on view and disparity zero-watermarks. Signal Process. **29**(1), 167–176 (2014)

89. A. Maalouf, M.-C. Larabi, C. Fernandez-Maloigne, A grouplet-based reduced reference image quality assessment. In: Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop On, IEEE pp. 59–63 (2009)

90. M. Zhang, W. Xue, X. Mou, Reduced reference image quality assessment based on statistics of edge. In: Digital Photography VII, vol. 7876, p. 787611 (2011). International Society for Optics and Photonics

91. J. Galbally, S. Marcel, J. Fierrez, Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. IEEE Trans. Image Process. **23**(2), 710–724 (2014)

92. Z. Wang, E.P. Simoncelli, Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. Hum. Vision Electr. Imaging **5666**, 149–159 (2005)

93. R. Soundararajan, A.C. Bovik, Video quality assessment by reduced reference spatio-temporal entropic differencing. IEEE Trans. Circ. Syst. Video Technol. **23**(4), 684–694 (2013)

94. L. Ma, S. Li, K.N. Ngan, Reduced-reference image quality assessment via intra-and inter-subband statistical characteristics in reorganized dct domain. In: Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conference (2011)

95. M.J. Scott, S.C. Guntuku, Y. Huan, W. Lin, G. Ghinea, Modelling human factors in perceptual multimedia quality: On the role of personality and culture. In: Proceedings of the 23rd ACM International Conference on Multimedia, ACM pp. 481–490 (2015)

96. X. Min, K. Gu, G. Zhai, M. Hu, X. Yang, Saliency-induced reduced-reference quality index for natural scene and screen content images. Signal Process. **145**, 127–136 (2018)

97. M. Ibrar-ul-Haque, M. Tahir Qadri, N. Siddiqui, Reduced reference blockiness and blurriness meter for image quality assessment. Imaging Sci. J. **63**(5), 296–302 (2015)

98. M.A. Saad, A.C. Bovik, C. Charrier, Blind image quality assessment: A natural scene statistics approach in the dct domain. IEEE Trans. Image Process. **21**(8), 3339–3352 (2012)

99. L. Ma, X. Wang, Q. Liu, K.N. Ngan, Reorganized dct-based image representation for reduced reference stereoscopic image quality assessment. Neurocomputing **215**, 21–31 (2016)

100. M. Carnec, P. Le Callet, D. Barba, Objective quality assessment of color images based on a generic perceptual reduced reference. Signal Process. **23**(4), 239–256 (2008)

101. L. Ma, S. Li, K.N. Ngan, Reduced-reference image quality assessment in reorganized dct domain. Signal Process. **28**(8), 884–902 (2013)

102. P. Le Callet, C. Viard-Gaudin, D. Barba, Continuous quality assessment of mpeg2 video with reduced reference. In: First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Phoenix (2005)

103. S. Wolf, M. Pinson, Video quality measurement techniques. 2002. (2002)

104. M. Narwaria, W. Lin, I.V. McLoughlin, S. Emmanuel, L.-T. Chia, Fourier transform-based scalable image quality measure. IEEE Trans. Image Process. **21**(8), 3364–3377 (2012)

105. Y. Arai, T. Agui, M. Nakajima, A fast dct-sq scheme for images. IEICE Trans. (1976-1990) **71**(11), 1095–1097 (1988)

106. L. Ma, S. Li, F. Zhang, K.N. Ngan, A reduced-reference perceptual quality metric for in-service image quality assessment. IEEE Trans. Multimedia **13**(4), 824–829 (2011)

107. D. Tao, X. Li, W. Lu, X. Gao, Reduced-reference iqa in contourlet domain. IEEE Trans. Syst. Man Cybern. B (Cybernetics) **39**(6), 1623–1627 (2009)

108. J.D. Valentich, Morphological similarities between the dog kidney cell line mdck and the mammalian cortical collecting tubule. Ann. N. Y. Acad. Sci. **372**(1), 384–405 (1981)

109. M.N. Do, M. Vetterli, The contourlet transform: An efficient directional multiresolution image representation. IEEE Trans. Image Process. **14**(12), 2091–2106 (2005)

110. X. Wang, G. Jiang, M. Yu, Reduced reference image quality assessment based on contourlet domain and natural image statistics. In: Image and Graphics, 2009. ICIG'09. Fifth International Conference On, IEEE pp. 45–50 (2009)

111. K. Rao, Yip,"Discrete Cosine Transform, Algorithm, Advantage and Applications" (Academic, New York, 1990)

112. F. Halsall, *Introduction to Data Communications and Computer Networks* (Addison-Wesley, Boston, 1985)

113. F. Yang, S. Wan, Bitstream-based quality assessment for networked video: A review. IEEE Commun. Mag. **50**(11), 1 (2012)

114. S. Decherchi, P. Gastaldo, R. Zunino, E. Cambria, J. Redi, Circular-elm for the reduced-reference assessment of perceived image quality. Neurocomputing **102**, 78–89 (2013)

115. K. Gu, V. Jakhetiya, J.-F. Qiao, X. Li, W. Lin, D. Thalmann, Model-based referenceless quality metric of 3d synthesized images using local image description. IEEE Trans. Image Process. **27**(1), 394–405 (2018)

116. K. Chono, Y.-C. Lin, D. Varodayan, Y. Miyamoto, B. Girod, Reduced-reference image quality assessment using distributed source coding. In: Multimedia and Expo, 2008 IEEE International Conference On, IEEE pp. 609–612 (2008)

117. S. Wolf, M.H. Pinson, Low bandwidth reduced reference video quality monitoring system. In: First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Arizona pp. 23–25 (2005)

118. I. Recommendations, Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference. ITU-R BT **1683** (2004)

119. M. Dye, R. McDonald, A. Rufi, *Network Fundamentals* (Cisco press, CCNA Exploration Companion Guide, 2007)

120. U. Engelke, M. Kusuma, H.-J. Zepernick, M. Caldera, Reduced-reference metric design for objective perceptual quality assessment in wireless imaging. Signal Process. **24**(7), 525–547 (2009)

121. T.S. Sector, *Opinion Model for Video-Telephony Applications* (Telecommunication Standardization Sec-tor, Geneva, 2007)

122. L. Ma, S. Li, K.N. Ngan, Reduced-reference video quality assessment of compressed video sequences. IEEE Trans. Circ. Syst. Video Technol. **22**(10), 1441–1456 (2012)

123. I.P. Gunawan, M. Ghanbari, Efficient reduced-reference video quality meter. IEEE Trans. Broadcast. **54**(3), 669–679 (2008)

124. M.H. Pinson, S. Wolf, An objective method for combining multiple subjective data sets. In: VCIP, pp. 583–592 (2003)

125. M.A. Usman, S.Y. Shin, M. Shahid, B. Lövström, A no reference video quality metric based on jerkiness estimation focusing on multiple frame freezing in video streaming. IETE Tech. Rev. **34**(3), 309–320 (2017)

126. Z. Wang, H.R. Sheikh, A.C. Bovik, No-reference perceptual quality assessment of jpeg compressed images. In: Image Processing. 2002. Proceedings. 2002 International Conference On, IEEE vol. 1, p. (2002)

127. T. Yamada, Y. Miyamoto, M. Serizawa, End-user video-quality estimation based on a reduced-reference model employing activity-difference for iptv services. In: Consumer Electronics, 2009. ICCE'09. Digest of Technical Papers International Conference On, IEEE pp. 1–2 (2009)

128. ITU-T RECOMMENDATION, P.: Subjective video quality assessment methods for multimedia applications (1999)

129. S. Nagata, Y. Ofuji, K. Higuchi, M. Sawahashi, Optimum resource block bandwidth for frequency domain channel-dependent scheduling in evolved utra downlink ofdm radio access. In: Vehicular Technology Conference, IEEE, 2006. VTC 2006-Spring. IEEE 63rd, vol. 1, pp. 206–210 (2006)

130. B. Girod, What's wrong with mean-squared error. Digital images and human vision, 207–220 (1993)

131. A. Albonico, G. Valenzise, M. Naccari, M. Tagliasacchi, S. Tubaro, A reduced-reference video structural similarity met-ric based on no-reference estimation of channel-induced distortion. In: Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference On, IEEE pp. 1857–1860 (2009)

132. G. Valenzise, M. Naccari, M. Tagliasacchi, S. Tubaro, Reduced-reference estimation of channel-induced video distor-tion using distributed source coding. In: Proceedings of the 16th ACM International Conference on Multimedia, ACM pp. 773–776 (2008)

133. Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment. IEEE Trans. Image Process. **20**(5), 1185–1198 (2011)

134. S. Wenger, Error patterns for internet experiments. ITU-T SG16 Doc. Q15-I-16r1 (1999)

135. T. Yamada, Y. Miyamoto, M. Serizawa, H. Harasaki, Reduced-reference based video quality-metrics using represent-ative-luminance values. In: Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, pp. 1–4 (2007)

136. T. Wiegand, G.J. Sullivan, G. Bjontegaard, A. Luthra, Overview of the h. 264/avc video coding standard. IEEE Trans. Circ. Syst. Video Technol. **13**(7), 560–576 (2003)

137. M. Rohani, A.N. Avanaki, S. Nader-Esfahani, M. Bashirpour, A reduced reference video quality assessment method based on the human motion perception. In: Telecommunications (IST), 2010 5th International Symposium On, IEEE pp. 831–835 (2010)

138. Z. Xiong, K. Ramchandran, M.T. Orchard, Y.-Q. Zhang, A comparative study of dct-and wavelet-based image cod-ing. IEEE Trans. Circ. Syst. Video Technol. **9**(5), 692–695 (1999)

139. H. Kim, J. Park, Efficient video quality assessment for on-demand video transcoding using intensity variation analy-sis. J. Supercomput. **75**, 1–15 (2018)

140. X. Wang, Q. Liu, R. Wang, Z. Chen, Natural image statistics based 3d reduced reference image quality assessment in contourlet domain. Neurocomputing **151**, 683–691 (2015)

141. T.-S. Wang, X.-B. Gao, W. Lu, G.-D. Li, A new method for reduced-reference image quality assessment. Journal of Xidian University **35**(1), 101–109 (2008)

142. Y.-H. Lin, J.-L. Wu, Quality assessment of stereoscopic 3d image compression by binocular integration behaviors. IEEE Trans. Image Process. **23**(4), 1527–1542 (2014)

143. C.T. Hewage, M.G. Martini, Reduced-reference quality assessment for 3d video compression and transmission. IEEE Trans. Consumer Electr. **57**(3), 1 (2011)

144. A. Mittal, A.K. Moorthy, J. Ghosh, A.C. Bovik, Algorithmic assessment of 3d quality of experience for images and videos. In: Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE, pp. 338–343 (2011)

145. C.T. Hewage, M.G. Martini, Reduced-reference quality metric for 3d depth map transmission. In: 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), IEEE, 2010, pp. 1–4 (2010)

146. C.T. Hewage, M.G. Martini, Edge-based reduced-reference quality metric for 3-d video compression and transmis-sion. IEEE J. Select. Top. Signal Process. **6**(5), 471–482 (2012)

147. A. Takahashi, D. Hands, V. Barriac, Standardization activities in the itu for a qoe assessment of iptv. IEEE Commun. Mag. **46**(2), 78–84 (2008)

148. G. Yang, D. Li, F. Lu, Y. Liao, W. Yang, Rvsim: A feature similarity method for full-reference image quality assessment. EURASIP J. Image Video Process. **2018**(1), 6 (2018)

Publisher's Note