Reprinted from
Journal of Imaging Science and Technology® 60(6): 060406-1–060406-15, 2016.
https://doi.org/10.2352/ISSN.2470-1173.2017.12.IQSP-228

Knowledge-based Taxonomic Scheme for Full-Reference Objective Image Quality Measurement Models

Atidel Lahoulou
University of Jijel, Computer Science Department, Jijel, 18000, Algeria

Mohamed Chaker Larabi
Université de Poitiers, XLIM SIC Department, Poitiers, 86962, France
E-mail: chaker.larabi@univ-poitiers.fr

Azeddine Beghdadi and Emmanuel Viennet
Université Paris 13, Sorbonne Paris Cité, L2TI, Villetaneuse, 93430, France

Ahmed Bouridane
Northumbria University, School of Comp., Eng. and Inf. Sciences, Newcastle upon Tyne NE1 8ST, United Kingdom

Abstract. *In this article, we propose a knowledge-based taxonomic scheme of the objective image quality assessment metrics including the key concepts involved for each approach. Our classification is constructed according to six criteria based on the information available at each stage of the design process. The novelty of the present classification scheme is that the six layers are linked via a single concept where each layer represents a single type of knowledge about: 1) the reference image, 2) the degradation type, 3) the visual perception field, 4) the human visual physiology and psychophysical mechanisms, 5) the processes of the visual information analysis, and finally 6) knowledge about perceptual image representation and coding. The first layer helps delineate boundaries between full-reference (FR) image quality assessment metrics, that are further classified through layers 2–6, and other families (reduced-reference [RR] and no-reference [NR]). In addition, gradual degrees are considered for knowledge about specific areas related to visual quality evaluation processes. The proposed taxonomic framework is intended to be stepwise, to help sorting out the fundamental ideas behind the development of objective image quality metrics often working on the luminance channel or marginally on the RGB channels. The aim is to congregate the already published classification schemes and to methodologically expand new aspects according to which an efficient and straightforward classification of the image quality assessment algorithms becomes possible. This is significant because of the increasing number of developed metrics. Furthermore, a systematic summarization is necessary in order to facilitate the research and application of image quality techniques. © 2016 Society for Imaging Science and Technology.*

INTRODUCTION

The evaluation of visual data quality is a critical task due to the new ways of consuming multimedia contents and the rapid growth of their availability. Indeed, there is a wealth of research on both subjective and objective image quality assessment (IQA) measures. Their goal is to provide computational models that can automatically and reliably predict the perceived quality of images across different scenes and distortion types. In other words, the predicted scores should be as close as possible to those that would be given by an average human observer. The latter task was the purpose of the Video Quality Experts Group (VQEG)[1] that validates and establishes subjective and objective standard approaches for visual data quality measurement.

Therefore, challenges behind developing objective image quality metrics (IQMs) are manifold. First, they can be employed for image quality monitoring during image acquisition, transmission and reproduction. Second, they can be deployed for benchmarking image processing algorithms designated for restoration and enhancement. Third, they can be embedded in compression and communication systems for parameter optimization.[2–4] In fact, an increasing range of applications and fields need automatic evaluation of the image quality either because of the impossibility to ensure human inspection or because of the plethoric amount of data. Indeed, the number of images and video hours uploaded/downloaded every minute to/from the web is just elusive. Furthermore, for a content to be noticed among all the incoming ones, high visual quality as well as high creative skills are required; hence new tools and editors dedicated to online image processing are being proposed. In addition, new research efforts are directed toward the extension of search engines to exploit the deep layer of the web not yet publicly accessible. This latter is significantly bigger than the visible part of the web everybody is browsing nowadays.[5,6] On the other hand, the IQA issue can be dealt with in a different way. In some cases, IQMs are designed to estimate the performance of imaging systems that pertain to printing technology, optical systems, capture systems, image quality engineering, etc. In this article, we focus only on computational metrics that take into account the visual

Received July 4, 2016; accepted for publication Oct. 20, 2016; published online Dec. 7, 2016. Associate Editor: Vien Cheung.

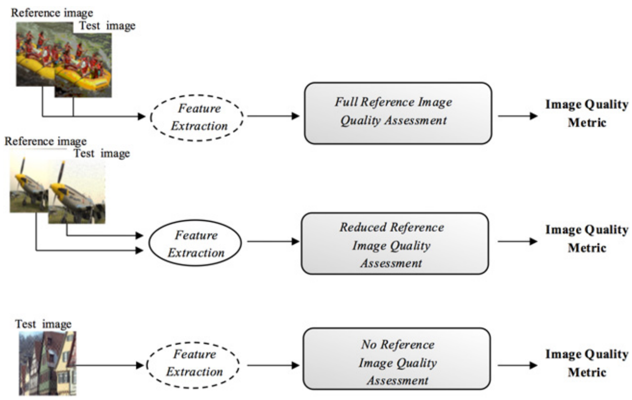

Figure 1. Relationship between subjective and objective image quality assessment.

perceived image quality, image fidelity or image distortion since they constitute the largest majority of quality metrics in the literature over the last 15 years.

Nevertheless, computational quality assessment metrics are usually benchmarked using subjective test results that constitute the ground truth. The aim of these experiments is to provide accurate, consistent and reliable predictions of the perceptual image appreciation. However, subjectively evaluating the quality of image content is an extremely difficult task due to the time and cost involved. Indeed, reliable subjective tests require the participation of a large number of human observers, evaluating the quality of images under restricted and controlled psychometric experimental conditions. In addition, perceptual quality may vary from one individual to another depending on observers' general experience (if he/she is expert in image processing or not), on their personal appreciation and may vary according to their mood. To alleviate this problem, the given individual ratings are aligned and averaged to compute the Mean Opinion Score (MOS). In some cases where reference images are also evaluated, the Difference Mean Opinion Score (DMOS) is derived instead of the MOS. More details on subjective ratings processing used to calculate the MOS/DMOS can be found in Ref.1.

Despite their drawbacks, subjective IQA measurements are essential to establish the perceptual predictive performance of objective models. As recommended by the VQEG,[1] this latter is defined in terms of several attributes namely correlation, consistency and monotonicity. Thus, Pearson's Correlation Coefficient (PCC) gives indication on linear correlation between objective and subjective ratings. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are computed to quantify the prediction accuracy. Finally, Spearman Rank Order Correlation Coefficient (RHO) and Kendall Rank Order Correlation Coefficient (TAU) serve to indicate the monotonicity measure. There exist an important number of image quality databases publicly available to the research community. They are systematically surveyed and analyzed such as in Refs. 2, 7 and 8. Figure 1 depicts the relationship between objective and subjective visual data quality evaluation approaches.

In addition, in order to strike a balance between precision and generality, we need to use some general terms so as to maintain the reader's attention on the core of our contribution rather than on the very complex terminology of image quality. Indeed, we use the term "quality" to refer to perceived visual quality, fidelity, similarity and distortion quantification. We also use the term "metric" to refer to many other different terms employed in the literature, such as model, measure, index, criterion and formula. For any discussion about image quality terminology, the reader is referred to Refs. 9 and 10.

In the present article, we propose a new knowledge-based taxonomy for the state-of-the-art objective FR quality metrics of grayscale still images. The purpose is more than just a summary or collection of IQA metrics. It is a hierarchical classification framework for understanding the relationships among the different categories of IQMs, as well as the connections between image quality on one side and knowledge about visual content and human visual system (HVS) on the other side. A knowledge structure of IQA metrics would provide a checklist to the image quality community. It is intended to help better see which research topics need to be deepened and more investigated in the future.

The article is organized as follows: section "Classification of FR image quality metrics (IQMs)" provides a critical overview of the prior classification schemes of image and video quality assessment metrics, the novelty of the proposed knowledge-based framework as well as a general description of this latter. Also, a brief outline on how to progress through the six layers of the proposed classification scheme is given. Each layer (from layer 1 to 6) representing a type of knowledge is then comprehensively described in section "Knowledge about the Environment" to section "Knowledge about Perceptual Image Representation and Coding". Finally, we conclude the article with discussion on the challenges, possible trends and key directions of the visual quality evaluation field.

CLASSIFICATION OF FULL-REFERENCE IMAGE QUALITY METRICS (IQMS)

In this section, we expose a new hierarchical classification framework for FR quality assessment metrics of monochromatic still images. There exists a long array of classifications of image and video quality assessment metrics in the literature, but only one taxonomic scheme introduced by Engeldrum,[9] as mentioned further. We then point out the shortcomings of the prior classification schemes and describe the new taxonomy proposed in this manuscript.

Prior Work

Given the rich literature of visual data quality assessment, it seems worthwhile to bring systematic summarization and comparison studies in order to facilitate the search and application of image quality techniques. Indeed, the very first classification of IQMs has been proposed in 2002 by Avcibas et al.[11] who divided the metrics into pixel-difference-based, correlation-based, edge-based, spectral-based, context-based and HVS-based measures. A short taxonomy has been later defined in 2004[9] where Engeldrum distinguished "Beauty Contest" and "Detection/Recognition" models employed for perceived quality evaluation and for visual systems, devices and algorithms benchmarking, respectively. Each of the two categories is further split into what he called "Ness"-based and "Physically"-based image quality, which refers to what is currently known as subjective an objective IQA, respectively.

Since then, many other reviews and classifications of IQMs have been published,[3,4,12–30] in order to serve the same purpose. Most of these articles appeared after 2010 where researchers particularly focused on objective FR IQA metrics. Among these, Engelke and Zepernick made a survey of the perceptual-based image and video quality metrics with special emphasis on quality of service.[12] The proposed scheme reported on (a) subjective versus objective quality approaches, (b) psychophysical versus engineering approaches, and (c) reference-based classification. In a different fashion, Gao et al. spotlighted the available biologically inspired image quality models and divided them into bionics and engineering categories referred to as "bottom-up" and "black-box" approaches, respectively.[13] As for Lin and Kuo, a distinction has been outlined between model-based and signal-driven perceptual visual quality metrics.[14] Each of the classifications presented in Refs. 12–14 considers only two categories for the objective perceptual FR IQMs; which is too restrictive compared to the growing number of approaches.

Several proposals have been published later with the aim to give more thorough classification schemes that encompass a wider range of IQMs. Chandler and Hemami[15] divided them into three groups including (a) mathematical convenient metrics, (b) near-threshold psychophysics-based metrics where a frequency-based decomposition is employed in order to quantify the visual fidelity of distorted images, and (c) a third unnamed category that brings together metrics based on the premise that structural content of high quality images most closely matches that of their original version. Another literature review of subjective and objective image quality algorithms according to their followed strategies and used techniques can be found in Ref. 16. In the latter, Thung and Raveedran categorized IQMs into FR, RR and NR families. The FR family is, in turn, decomposed into mathematical metrics, HVS-based metrics and a third unnamed category of metrics. A similar classification has been introduced in Ref. 17 by Pedersen and Hardeberg who congregated the schemes presented in Refs. 15 and 16. They described four groups of metrics namely (a) mathematically based metrics, (b) low-level metrics, (c) high-level metrics and (d) a group of other metrics which either do not fit in the previous classes or use combined strategies. More recently, Chandler proposed a comprehensive survey dealing with full-, reduced- and free-reference IQMs.[4] Likewise, FR metrics are categorized into methods based on (a) HVS models, (b) image structure, (c) image statistics and machine learning, and (d) a general category based on other techniques. Although surveys given in Refs. 4, 15–17 provided extensive lists of references, their major weakness is related to the use of a general class that encloses all the quality metrics that do not fit the previously defined categories. This is due to the very large amount of concepts and methods used in the IQA framework, which makes it so hard to delineate boundaries between them. However, the use of a general category widely limits the classification scheme.

Other researchers defined more classes in their schemes as in Refs. 18 and 19. Indeed, Chikkerur et al. introduced a classification scheme for FR and RR objective video quality assessment models.[18] According to the suggested scheme, three preliminary classes are defined: (a) traditional point-based metrics, (b) natural visual characteristics divided into natural visual statistics and natural visual features classes and (c) perceptual models including frequency and pixel domain categories. A different categorization was suggested by Akramullah in Ref. 19 for subjective and objective video quality assessment. So, metrics are sorted according to the following categories: (a) error sensitivity, (b) structural similarity, (c) information fidelity, (d) spatio-temporal, (e) saliency and (f) network-based approaches. Both reviews in Refs. 18 and 19 give performance comparison of presented techniques, with the aim to demonstrate the differences between them. Nevertheless, it is still hard to categorize some models and techniques since there are bridge concepts involved in many image and video quality assessment metrics. This makes it unlikely to classify one metric in one class rather than in another since there are no links between the proposed classes.

Another interesting scheme was presented by Wang and Bovik in Ref. 3. The authors provided a classification of IQMs based on three types of knowledge about: (a) the reference image, (b) the distortion process and (c) the HVS. Seven families of models are then distinguished following this classification. They are all linked via a single notion, namely knowledge. Furthermore, the HVS is extremely complex and constitutes the main concept upon which all bio-inspired image and video quality approaches are built. The review aforementioned considers bottom-up and top-down approaches for the latest stage of the classification. However, knowledge about the HVS—when available—can neither be simply considered without taking into account the visual stimuli (image or video content, related coding and representation), nor without accounting the multiple aspects of the HVS (visual perception, related physiology and psychophysics, information analysis mechanisms, . . .). These notions need to be deepened and more detailed as it is attempted in the present article. In addition to the described works, one can cite other substantial surveys and summaries,[20,30] where the proposed categorization, when provided, is similar to discussed ones. Furthermore, some efforts have been dedicated to study and compare the predictive performances of the objective quality approaches.[14,17,18,28,31–33]

Motivation

The aforementioned surveys from the literature are undoubtedly very helpful and contain very high-level technical material. However, most of the proposed classifications are built upon a conceptual categorization approach, which is based on what IQA metrics do. This may reduce the task of visual data quality assessment to the testing of opponent techniques. We do believe that the IQA problem should be looked at as a "design process" where, according to the progress of other related research fields, a given amount of knowledge is available and may be efficiently exploited.

In this article, we develop a knowledge-based taxonomic scheme of an up-to-date array of objective FR IQA algorithms. It includes key relevant models from the conventional approaches to the most prominent ones, with a primary focus on computational metrics dedicated to the evaluation of distortion or fidelity, constituting the vast majority of IQMs in the literature. Hence, our aim is twofold: first, we describe the diversity of categories of metrics and then, we highlight the related types of information that have led to this diversity. However, only abbreviated descriptions of IQMs are provided for the classified IQMs since the emphasis is on underlying connections between the different categories rather on the technical aspects. Also, no analysis of the predictive performance and computational costs of the classified IQMs is performed.

The novelty of the present classification methodology lies in the fact of showing the whole framework of IQA algorithm development by defining a set of six layers for the design process. The layers are linked via a single concept in such a way that each layer represents a type of knowledge. In addition, gradual degrees are considered for knowledge about specific areas ranging from deep level to superficial level. The proposed taxonomic scheme is intended to be stepwise, to comprehensively review objective quality metrics without focusing on specific color fidelity measures and help sorting out the fundamental ideas behind their development. It is worth noting that the proposed scheme is not intended to include quality assessment metrics dedicated to video, audio, multimodal or any other type of content.

Description of the Proposed Knowledge-based Taxonomy

The proposed knowledge-based taxonomy is illustrated in Figure 2. It first classifies an IQM according to the prior knowledge about the reference image considered as the pristine version. We obtain FR, RR or NR families of models whether this information is complete, partial or unavailable at the quality prediction stage. The second layer considers the knowledge about the possible degradations which can be seen as very important information to be supplied about the environment. In this way, IQMs can be either designed for a specific distortion or general purpose according to whether or not information or assumptions about the nature of the distortion are exploited. This allows definition of the distortion-aware metrics and the fidelity metrics. Our classification scheme further subclassifies the distortion and fidelity metrics according to the knowledge about the visual perception into biologically inspired metrics (where this type of information is required) and traditional error-based ones (where no information about visual perception is used). As per the application level of the findings on visual physiology and cognitive phenomena, the biologically inspired techniques in the fourth layer can be considered either perceptual or signal-driven models, respectively. In turn, the category of perceptual models encompasses vision-model-based and embedded visual attention (VA) metrics depending on the degree of understanding of the visual information analysis mechanisms that can be deep or more or less superficial. As for the signal-driven class, it covers both natural scene statistics (NSS) and visual features-based metrics as well as machine learning oriented techniques. The last layer of our classification takes into account the utilized knowledge about perceptual image representation and coding which generates twenty different subclasses for the whole existing objective FR IQMs. In this sense, each IQM belonging to one of the ten subclasses of layer 6 of Fig. 2 (content-independent metrics, . . . mid-level properties-based metrics) can be either distortion-aware metric or fidelity metric, as they have been broken down in layer 2 of the taxonomy.

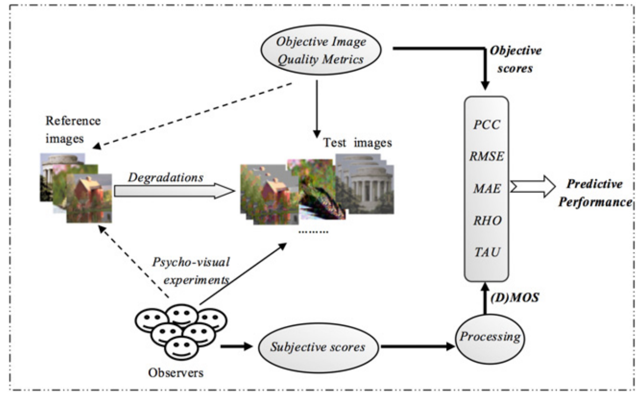

Figure 2. Overview of full-reference, reduced-reference and no-reference image quality assessment approaches.

KNOWLEDGE ABOUT THE ENVIRONMENT

In this section, two layers of the proposed scheme are described. They concern knowledge about the reference image and knowledge about possible degradations, respectively.

Knowledge about the Reference Image: Full-reference, Reduced-reference and No-reference IQMs

Depending on the amount of available information about the reference image at the time of estimation of the distorted version(s), it is rather common to distinguish three broad families of models in the literature as schematized in Figure 3 below. Note that the feature extraction step is optional for both FR and NR IQA families of models.

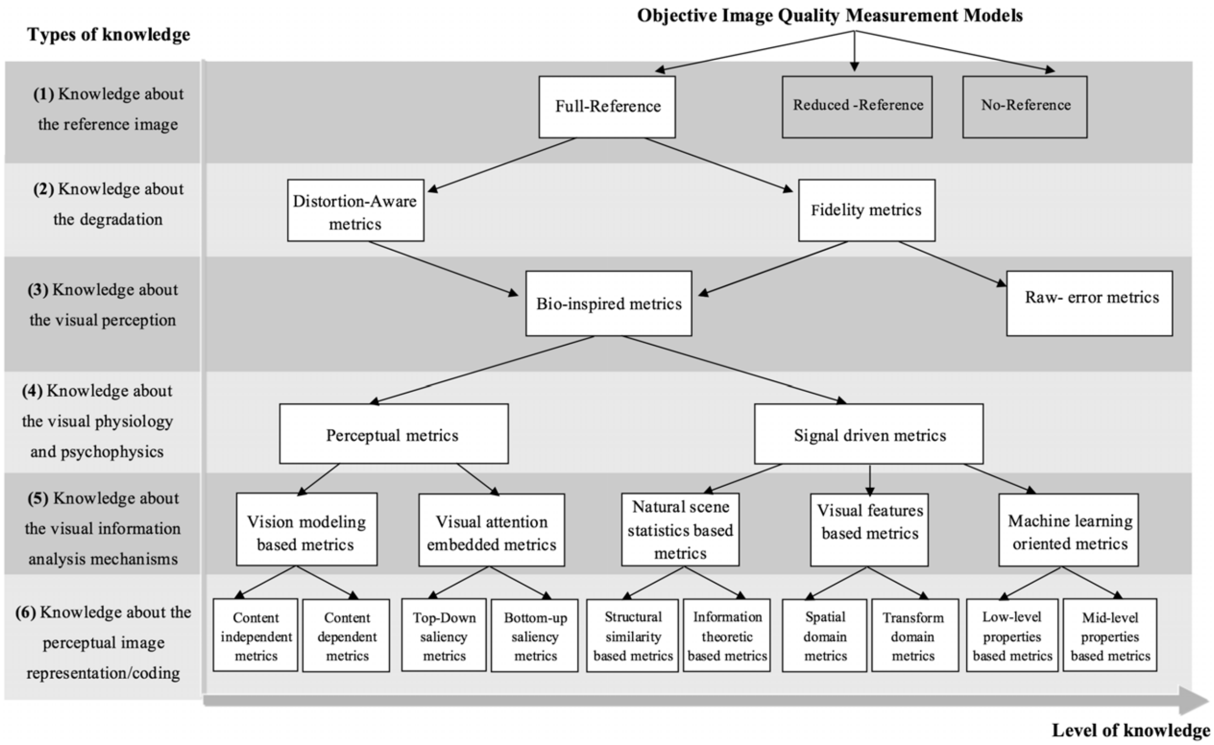

Figure 3. Knowledge-based taxonomic scheme for objective image quality measurement models. Each layer (1–6) represents the type of knowledge required to make the subclassification at each stage. The horizontal axis represents the level of knowledge ranging from deep level to superficial level of knowledge.

√ Full-reference models (FR)

Here the reference image is available when evaluating its impaired version(s). The task consists in a pairwise comparison and image quality evaluation can be seen in that case as an image fidelity problem. The prediction is expected to be fast and to correlate with human subjective appreciation. Some sets of image features can eventually be extracted prior to the image quality evaluation. When one of the features requires both reference and test images to be calculated, the quality metric falls in the FR category.

√ Reduced-reference models (RR)

In practice, it is often impossible to have the reference image for the evaluation of the distorted version(s) available. This is the case of broadcasting or streaming applications, to give a few examples where there is no access to the pristine content at the user side. Since the reference image is not provided, a feature vector giving relevant information can be transmitted with the aim to control the quality of transmitted visual data. Methods based on these features are fast, but their lack of genericity in addition to their variable performance make their use restricted to some specific applications.[2]

√ No-reference models (NR)

Also called "blind models," they attempt to evaluate the quality of an image without any cue about its reference. They are often distortion-oriented and are complicated to elaborate but are very useful and interesting for many applications. The feature extraction step here is optional.

Knowledge about the Degradation: Distortion-aware versus Fidelity IQMs

In this section, we outline the difference between distortion-aware and fidelity IQMs. Indeed, the *a priori* distortion nature and/or intensity an image undergoes may be available at the stage of appreciation of perceived quality. Thus an IQM can be considered as distortion-aware (also called specialized) or distortion-unaware (commonly called generic). Traditionally, the distortion-aware family of metrics requires knowledge or assumptions about the type of the distortions since the perceived quality may be affected by different artifacts. Research in this category has been conducted on designing algorithms that directly measure the possible impact created by specific image distortions. Most of the efforts have been directed toward NR and RR IQMs as well as video quality metrics, but very few are the works that have been interested in specialized FR quality metrics of still images. In addition to metrics early introduced in Refs. 34–36 but never confronted to subjective image quality evaluation, number of distortion-aware metrics have been developed and benchmarked. They are generally dedicated to noise measure and to coding applications, as discussed in section "Knowledge about Perceptual Image Representation and Coding".

On the other hand, generic approaches are either fidelity metrics, consisting of perceived image differences in the presence of both original and test images, or perceived quality metrics. Since our article deals only with FR metrics, then distortion-unaware category is referred to as fidelity one.

Furthermore, the real challenge for FR metrics is to be robust to wide panel of distortions. This property still remains very difficult since the number of degradations an image can be subjected to is somewhat unlimited. In addition, an image may undergo more than a single degradation at a time.

KNOWLEDGE ABOUT THE HUMAN VISUAL SYSTEM

In this section, three layers of the proposed scheme that concern knowledge about the HVS are introduced.

Knowledge about Visual Perception: Bio-inspired versus Raw-error IQMs

Fig. 2 shows that objective FR IQA algorithms can be, in turn, categorized on the basis of the knowledge about the outcomes of the visual perception field. Visual perception is the ability to interpret and understand the surrounding environment by processing information contained in the retinal images acquired by our eyes. This allows humans to distinguish objects, events, people and situations. Visual perception is very complex and requires some awareness coming from different fields such as psychology, cognitive science and neuroscience.

Raw-error-based measures do not rely on any information about human vision while biologically inspired quality metrics are built upon the characteristics or the functionalities of the HVS and the visual cortex. This makes them more consistent with the subjective judgment in spite of the fact that knowledge on the visual process is still incomplete.[37]

√ Biologically inspired IQMs

Considerable efforts have been devoted to the development of new objective quality metrics accounting for human vision characteristics. A variety of models belonging to this category exists and can be classified using our scheme according to the knowledge about the HVS anatomy and the psychophysical mechanisms. Hence, the biologically inspired metrics can be split into perceptual and signal-driven categories as shown in section "Knowledge about Physiology and Psychophysical Phenomena: Perceptual versus Signal Driven IQMs".

√ Raw-error-based IQMs

Initial investigations on objective IQA focused, for decades, on mathematical metrics based on error quantification between a pair of images (original and impaired). Among this category, signal to noise ratio (SNR), peak signal to noise ratio (PSNR) and mean squared error (MSE) were derived by hypothetically considering that the image distortion is produced by only additive noise, which is independent from the signal. However, the latter metrics do not rely on any model of noise. They can be seen as the precursors and are still widely used especially for compression evaluation purposes. Their popularity is due to their simplicity and their very low computational cost, at the price of a lack of correlation with human judgment. More discussion about the predictive performance of the raw-error metrics can be found in Refs. 33, 38–40.

Knowledge about Physiology and Psychophysical Phenomena: Perceptual Versus Signal-driven IQMs

The HVS is very complex and not fully understood yet. Most visual properties of the HVS are not intuitive, hence the development of physiological and psychophysical experiments that have been conducted to understand the involved phenomena. Psychophysics involves the study of the response (psycho) to a known stimulus (physics) in order to establish an empirical relationship between them.[13,37]

As it appears from Fig. 2, two different approaches can be discerned to formulate the biologically inspired image quality evaluation problems according to whether we make use of the research findings of the vision community or just make assumptions/approximations of the functionalities. On the one hand, perceptual algorithms' design requires scrupulous understanding of the HVS as well as the inherent psychophysical features, such as Contrast Sensitivity Functions (CSF), luminance and contrast masking phenomena, perceptual decomposition into channels, Just Noticeable Difference (JND), saliency and VA. Thus, the HVS functional components that may be relevant to the process of image quality appreciation are simulated, combined and integrated into the quality prediction schemes.

In contrast, the signal-driven approach leads to establish visual perception behavioral representations by making hypotheses about the overall functionalities of the HVS viewed as a black box. Systems that belong to this family simulate the HVS feedback when evaluating the visual quality of an image but do not necessarily operate in the same manner. This difference does not matter provided that the quality measure successfully predicts the human judgments.[3]

Knowledge about Visual Information Analysis Mechanisms

Tremendous work continues to be devoted to develop IQMs with the constraint of being consistent with the human judgment of quality. Knowing that the quality appreciation activity is supported by the analytical reasoning abilities of human, one can understand why it may vary from one individual to another for the same visual data. This fact may be taken into account in order to distinguish the scenario-based (perceptual models) and the hypothesis-based (signal-driven methods) analytical techniques in terms of the knowledge about the mechanisms of visual information analysis.

√ Vision modeling versus VA for IQMs

Both perceptual and signal-driven IQMs are developed to mimic the cortical decomposition performed by the HVS. However, the first family of techniques requires in-depth understanding of the human visual mechanisms.

Vision modeling approaches attempt to scrupulously build an HVS model with regard to quality evaluation, whereas another recent trend (attention-based models) aims at integrating VA principles into the process of quality evaluation. Most of those models are based on the hypothesis that the image distortion in salient regions generates more annoyance for the observer. However, using saliency information in the quality function in order to improve its predictive performance remains an open issue since it depends on many parameters such as the features used for saliency prediction and/or the fusion of conspicuity maps.[41]

√ NSS, visual features and machine learning oriented IQMs

Unlike model-based and attention-based philosophies which consist in designing models related to "what" does the HVS achieve when shown a distorted image, the signal-driven image quality measures operate based on assumptions on "how" the overall human visual cortex respond to such stimuli. They consider the HVS as a black box and focus on (a) signal modeling using Natural Scene Statistics for example, (b) processing of visual signals under consideration using low-level image attributes (visual features) or (c) input–output matching (machine learning).

KNOWLEDGE ABOUT PERCEPTUAL IMAGE REPRESENTATION AND CODING

In the context of IQA, the content can undergo some transformations and features computation prior to the quality/impairment measurement when it is believed that this process is relevant to the subjective quality appreciation and may positively affect the predictive performance of the algorithm. This last criterion allows dividing each of the categories obtained in the previous stage into two subcategories as depicted in Fig. 2. It is thoroughly explained in the following subsections. We also provide examples, together with abbreviated descriptions, of metrics that belong to each defined class. The chosen IQMs examples have significant insights in the current state of the art and derive from various approaches considered in the classification. We also categorize the IQMs according to their respective classes as defined in taxonomic scheme of Fig. 2.

Vision Modeling-based Metrics: Content-independent versus Content-dependent IQMs

Emphasis of the content-independent model-based measures is placed on human visual properties including the well-known and deeply explored CSF, temporal/spatial multichannel decomposition, various masking effects, JND function and luminance adaptation. Earliest HVS-inspired image distortion measure (DM) has been developed by Mannos and Sakrison[42] who exploited, for the first time, vision science findings to the image processing field. They proposed a model of human CSF that is still quite popular and widely used in computational metrics. The rationale behind using such a function is to account for the nonuniform sensitivity of the HVS and to filter images as they were perceived by the human eyes. The DM is then given by the squared difference between the resulting filtered images.

Later, Faugeras[43] used the CSF to build the first color image DM. More elaborate FR impairment estimators mimicking the HVS have been researched where the basic idea is the multichannel decomposition drawn from Campbell and Robson's experiment.[44] In fact, several multiple subband vision models have been proposed using sophisticated transforms such as the cortex transform in Ref. 45 which was used further in a modified version to design the well-known Visible Difference Predictor (VDP) by Daly.[46] It provides a visibility map indicating the areas where two images differ from the human perception point of view. The Sarnoff Visual Discrimination Model (SVDM) was designed by Lubin 47 and uses a Laplacian pyramid to separate the spatial frequencies of the image. Safranek and Johnston[48] used the generalized quadrature mirror filter transform for decomposing the image signals into sixteen subbands of the frequency space. Following the same objective, Heeger and Teo introduced an IQM in Ref. 49 including a steerable pyramid decomposition with six orientations. The latter transform has also been used to perform the spatial decomposition of images. It resulted into the design of the Perceptual Distortion Metric (PDM) by Winkler.[50] After that, Gao et al.[51] proposed the incorporation of a technique based on multiscale geometric analysis. Though, the complexity of the multichannel decomposition imposes heavy computational and memory requirements, even for moderately sized images[15] compared to the single-channel approaches. These latter demonstrate a relative simplicity but a less predictive accuracy since they are not related to the neural responses in the primary visual cortex.

In order to meet predictive performance requirements of IQA algorithms as well as computational efficiency, simple subband decomposition transforms have been tested giving satisfactory results. Some examples include metrics using the Discrete Cosine Transform (DCT) such as in the DCT-based Perceptual Error Measurement (DCT-PEM) proposed by Watson in Ref. 52 Moreover, Watson et al.[53] proposed a method that can quantify noise by means of the wavelet transform. Lu et al. introduced a metric in Ref. 54 based on wavelet transform that takes into account the variance of the visual sensitive coefficients in order to measure the quality of a distorted image. The wavelet version of the Daly's VDP has been suggested by Bradley in Ref. 55 under the name of Wavelet Visible Difference Predictor (WVDP). The Haar wavelet has been later employed by Lai and Kuo in Ref. 56 to implement an IQM for compressed images. On the other hand, DCTune has been proposed by Watson in 1993.[57] It is a tool originally developed for the visual optimization of DCT quantization matrixes in the JPEG compression scheme. It is based on the "image-dependent perceptual (IDP)" method to solve the bitrate–perceptual quality tradeoff. IDP method implements the following properties of the HVS: luminance masking, contrast masking, perceptual error and JNDs, spatial error pooling and frequency error pooling. DCTune has been commonly used within the image quality community as an IQM.

The interaction of the eye with image content is therefore ill-considered by the first family of image quality measures based on vision modeling. Humans instinctively analyze the visual content whether they are asked to evaluate its perceived quality or to perform a different test. Hence, in the content-dependent IQA methods, HVS functionalities are underlined rather than the HVS properties. The existing models describe color vision, frequency-orientation analysis, contour detection, perceptual and localization of patterns, object discrimination, and visual memory. Criterion 4 (C4) introduced by Carnec et al. is a quality estimator falling in this category.[58] It assesses perceived quality using elaborate models of several processing areas of the visual cortex as shown earlier but also by extracting structural information from the representation of images in a perceptual space. Albeit the C4 criterion is based on similarity measures, the authors consider it as a RR IQM since it uses features as side information that can be transmitted.

Visual Attention Integrated Methods: Top-down versus Bottom-up Models

Approaches consisting in the integration of VA into IQM —with the aim to potentially improve their predictive performance—have been built upon experimental results showing that the image distortion in salient regions is more annoying for human observers.[59] VA models are incorporated in a number of other applications such as object detection and recognition, image retrieval, image retargeting and image/video coding.[60] An overview of the most significant saliency models in the literature is available in Ref. 61. Computational VA models can be divided into two groups: top-down models driven by visual processing tasks and bottom-up models driven by low-level stimuli. From the quality assessment point of view, the top-down family of metrics is voluntary influenced by high-level features such as semantic objects in the scene (faces, bodies,. . .) or their location, while in the bottom-up category, it is supposed that VA is driven by low-level information of the scene (intensity, color, orientation,. . .).[60] Several studies have been conducted to evaluate the potential benefits brought by incorporating VA into already existing and commonly used IQA metrics[62–71] by using bottom-up or top-down saliency models, separately or jointly.

In addition to these investigations, novel algorithms have been developed in the aforementioned framework. Sadaka et al.[72] integrated the famous bottom-up Itti's model[73] into a NR sharpness metric using a multiplicative weighting function. This metric is stated here because it was the first original image quality measure in the VA-driven class. Later on, Zhang et al. proposed a new bottom-up saliency model named Saliency Detection by combining Simple Priors (SDSP)[74] used later to develop the Visual Saliency-Induced Index (VSI).[75] The visual saliency map is used as a weighting function to characterize the importance of local image regions. Top-down saliency models have especially been used for the design of NR perceptual quality measures. To the best of our knowledge, one original algorithm can be found for the FR metrics where Saha and Wu designed the Global Local Distortion using the Spectral Residual- based saliency metric (GLD-SR) and the Global Local Distortion using the Phase Fourier Transform-based saliency metric (GLD-PFT).[76]

After the effervescence known by this field, it has been concluded in Ref. 59 that the VA embedding is generally positive; however, it is more significant in video than in image applications. This is probably due to the fact that the new VA models have been tested onto already existing image quality measures especially because of their computational efficiency and competitive accuracy. It is consistent to think that if new metrics are developed by embedding a VA strategy, the results would be more successful as in the case of VSI introduced in Ref. 75.

Natural Scene Statistics: Structural Similarity versus Information Theoretic Based IQMs

The class of FR IQMs based on statistical features can be subdivided into structural similarity and information theoretic methods according to whether the concern is focused on image fidelity or on information fidelity, respectively.

One of the findings on the HVS is that it is highly adapted to extract structural information from images. By following the assumption that image quality degradation is due to the loss of structured information in images, then designing a quality metric that measures structural distortions should have good correlation with the perceived image dissimilarity.[77]

In Ref. 77, Wang et al. suggested that the HVS behavior can be decomposed into three independent channels: Luminance, Contrast and Structure. The universal quality index (UQI) has then been constructed using the product of the comparison equations of the three image components pairs. The formula incorporates the mean pixel intensity values, the standard deviation of the pixel intensity values and the covariance between pixel intensity values of the reference and the test images.

It has been found later that the UQI suffers from an instability problem. One solution to solve the latter problem has been suggested in Ref. 78 by making some modifications to the luminance, contrast and structure comparison definitions. The changes consist of introducing positive nonzero parameters (α, β, γ) to define the importance of each of the three components, which has led to a new quality metric called structural similarity index (SSIM) that uses the same image attributes as the UQI.

Three variants of the SSIM have been suggested later including the multiscale SSIM (MS-SSIM)[79] which incorporates the variations of the image resolution and viewing conditions, the SSIM with automatic down-sampling (MSSIM),[80] and the Complex Wavelets-SSIM based on the principle that structural information is more contained in the phase than in the magnitude of the signal.[81]

Shnayderman et al.[82] proposed a FR quality metric for multichannel image based on Singular Value Decomposition. The M-SVD measures the distortion as a function of distance between the reference and test image block singular values. A global measure is then derived from the individual DMs of all images' block pairs. Thus, the global error is the average of the differences between the distance measure for each block and the median of all block distance measures. Based on the same theorem, the R-SVD quality predictor has been developed later by Mansouri et al.[83] where the right singular vector matrix of the original image is used.

Since 2004, a new approach to IQA problems has emerged and new quality metrics have been developed using an information and communication theoretic framework. The information fidelity criterion (IFC)[84] and its extensions visual information fidelity (VIF) index and pixel-based VIF[85] metrics proposed by Sheikh et al. belong to a different class of IQA methods, built upon NSS models. The employed premise in the present case is that visual fidelity can be accurately quantified if it is known how much Shannon information the test image brings about from its reference version. Following Shannon's communication scheme, the transmitter, the channel and the receiver correspond to the reference source image, the distortion model applied on it and the test generated distorted image, respectively.

An example of specialized IQMs belonging to the NSS class is the local standard deviation-based image quality (LSDBIQ). It has been developed by Gore et al.[86] to assess the quality of JPEG-coded images. The approach is based on the comparison of the local standard deviation of the original and test images.

Visual Features Based Methods: Spatial Domain versus Frequency Domain Based IQMs

Metrics based on natural visual statistics principles discussed in section "Natural Scene Statistics: Structural Similarity versus Information Theoretic based IQMs" are particularly attractive due to their mathematical foundations which facilitate their analysis and optimization. However, because these metrics do not consider the detectability of distortions, their applicability to determine whether or not a distorted image is of perfect visual quality remains unclear.[15] The idea to overcome this drawback is to use some well-known signal processing techniques that have similar features as the human visual perception. Image quality metrics relying on the principles of characterizing low-level but also mid-level visual properties are reported in this subsection.

A new era has been initiated by the development of image quality measures that incorporate global HVS properties to simulate the perceived reference and test images. The Picture Quality Scale (PQS) for achromatic image coding[87] and the Noise Quality Metric (NQM) for image restoration purposes[88] are typical examples. Both PQS and NQM belong to the distortion-aware IQMs class. PQS measures the image quality degradation caused by coding impairments. It is a linear combination of three weighted factors of distortion, including the amount of error, the location of error and the structure of error. NQM actually includes two quality measures: a DM (DM) computed to quantify the effect of linear frequency distortion, and a NQM computed to measure the effect of additive noise.

From the beginning of the 90s, one can notice a revival of the raw mathematical metrics including SNR and PSNR combined to some basic human visual features in the state of the art of image quality evaluation. This has resulted in WSNR, $\mathrm{SNR}_{\mathrm{WAV1}}$, $\mathrm{SNR}_{\mathrm{WAV2}}$, SNR_W, PSNR_W, PSNR-HVS, and PSNR-HVS-M explained below.

The weighted version of the signal to noise ratio (WSNR) has been derived by Mitsa et al.[89] using the spatial Contrast Sensitivity Function (CSF) defined as the ratio of the averaged weighted signal power to the average weighted noise power. The HVS-based PSNR (PSNR-HVS) developed in Ref. 90 takes only the CSF into account as a visual feature.

The PSNR-HVS-M metric is a frequency-based version of the PSNR-HVS where an improvement has been brought thereafter by introducing the model of visual correlation between-coefficient contrast masking of DCT basis functions based on the HVS.[91] Iordache et al.[92] developed an image dissimilarity measure based on a joint spatial/spatial-frequency representation using Wigner–Ville distribution. The SNR_W measure is built upon the assumption that structured distortions are more annoying than unstructured distortions. $\mathrm{SNR}_{\mathrm{WAV}}$ is another image DM proposed by Beghdadi et al.[93] and based on nonredundant wavelet decomposition. The multiresolution analysis computed by means of wavelet transform allows accounting for the effect of the distortions at different scales. Two families of wavelets have been employed, the biorthogonal 9/7 wavelets and the cubic spline wavelets giving birth to the $\mathrm{SNR}_{\mathrm{WAV1}}$ and $\mathrm{SNR}_{\mathrm{WAV2}}$ measures, respectively. PSNR versions of the aforementioned image quality measures including PSNR_W and $\mathrm{PSNR}_{\mathrm{WAV}}$ can be found in Ref. 94.

Another quality estimator, the Visual SNR (VSNR) proposed by Chandler et al. in Ref. 15 evaluates image quality according to a contrast model accounting for low-level HVS properties and mid-level HVS property of global precedence.[95] In Ref. 96, Larson and Chandler proposed an IQA metric they named most apparent distortion (MAD). MAD combines two different adaptive strategies of the HVS according to the amount of distortion in images: a detection-based and an appearance-based perceived distortion strategy. The first strategy is adopted by the HVS when observed images are high quality. It is modeled via local luminance and contrast masking. In the case of low-quality images where the distortion is most apparent, changes in the local statistics of log-Gabor coefficients are employed to model the second HVS strategy.

Another example of distortion-aware IQMs is DCTex proposed by Zhang et al.[97] DCTex metric is dedicated for coded images with the key assumption that the signal error in each subband and each local region contributes to the entire distortion independently. Therefore, the distortion is decomposed into independent blocks and subbands by taking into account two properties of the HVS, including the texture masking effect and contrast sensitivity function (CSF).

The Riesz-transform-based feature similarity metric (RFSIM)[98] was proposed based on the hypothesis that the HVS perceives an image mainly according to its low-level features at key locations. The 1st and 2nd order Riesz-transform coefficients of the image are taken as image features. Moreover, key locations are indicated by a feature mask generated by the Canny edge detection. The similarity index between the reference and test images is then derived by taking only the Riesz features coefficients within the feature mask.

Another feature similarity index for FR images (FSIM) has been designed by Zhang et al.[99] It is based on the finding that image understanding is essentially based on the low-level features of image. In the first stage, the image local quality is characterized by combining the phase congruency (PC) and the image gradient magnitude (GM) used as primary and secondary features, respectively. In the second stage, the PC is used again as a weighting function to derive an overall quality score from the previous similarity map.

Machine Learning Oriented Metrics: Low-level versus Mid-level Properties Based IQMs

According to the machine learning oriented approach, the visual data quality evaluation is generally formulated as a supervised regression problem (also called function approximation) where the goal is to train the system to find the best input–output relationship by optimizing the difference between the estimated output and the desired one. In this application, the output is the subjective quality rating while the input is the image for instance. Hence, the input vector is as large as the image size, which causes a complex phenomenon named "curse of dimensionality." One possible solution that attempts to overcome this pitfall is the use of feature extraction and selection.[3] Furthermore, finding appropriate image features is critical for the machine learning oriented image quality measures, which are mostly artificial neural network based.

Many methods extract low-level descriptors. For instance, in Refs. 100, 101, the authors estimate image quality using a circular back-propagation neural network using two different set of features. In Ref. 102, Bouzerdoum et al. used a Multilayer Perceptron (MLP) to predict the MOS. The results were obtained by feeding the MLP with statistical indicators vector including the mean, the standard deviation, the covariance and the mean square error of both reference and test images. Narwaria and Lin[103] proposed an IQA algorithm based on support vector regression. The input features are the singular vectors out of singular value decomposition. The machine learning techniques have also been used for feature selection to define the most relevant image descriptors to image quality evaluation task. Support Vector Machines have been employed in Refs. 104, 105. The authors have further used the multilayer perceptron in Ref. 106 for feature selection on the same set of low-level image features mentioned above. After identifying the image attributes that are the most relevant to the image quality appreciation, Lahoulou et al.[107] derived two measures called Error-based Cost Function (ECF) image quality measure and Correlation-based Cost Function (CCF) image quality measure designed to optimize the accuracy

Table I. Summary of bio-inspired full-reference image quality metrics belonging to vision modeling-based methods from perceptual category, where (DM = Distortion-aware metrics, FM = Fidelity metrics).

Vision Modeling-based Metrics						
Metric's description	Symbol	DM	FM	Authors	Year	Ref.
Content Independent Metrics						
Distortion measure for monochrome still images	—	x		Mannos et al.	1974	42
Generalized quadrature mirror filter transform-based metric	—		x	Safranek et al.	1989	48
Visible Difference Predictor	VDP		x	Daly	1993	46
DCT-based Perceptual Error Measurement	DCT-PEM		x	Watson	1993	52
Visual optimization of DCT quantization matrixes	DCTune	x		Watson	1993	57
Sarnoff Visual Discrimination Model	SVDM		x	Lubin	1995	47
Steerable pyramid decomposition-based measure	—		x	Heeger et al.	1995	49
Noise quantization measure using DWT	—	x		Watson et al.	1997	53
Perceptual Distortion Metric	PDM	x		Winkler	1998	50
Wavelet Visible Difference Predictor	WVDP		x	Bradley	1999	55
IQM based on the variance of DWT coefficients	—		x	Lu et al.	2008	54
Multiscale geometric analysis-based measure	—		x	Gao et al.	2009	51
Content Dependent						
Criterion 4	C4		x	Carnec et al.	2008	58

Table II. Summary of bio-inspired full-reference image quality metrics belonging to VA embedded methods from perceptual category, where (DM = Distortion-aware metrics, FM = Fidelity metrics).

Visual Attention Embedded Metrics						
Metric's description	Symbol	DM	FM	Authors	Year	Ref.
Bottom-up saliency						
Saliency Detection by combining Simple Priors	SDSP		x	Zhang et al.	2013	74
Visual Saliency-Induced Index	VSI		x	Zhang et al.	2014	75
Top-down saliency						
Global Local Distortion using the Spectral Residual- based saliency	GLD-SR		x	Saha et al.	2015	76
Global Local Distortion using the Phase Fourier Transform saliency	GLD-PFT		x	Saha et al.	2015	76

Table III. Summary of bio-inspired full-reference image quality metrics belonging to NSS-based methods from signal-driven category, where (DM = Distortion-aware metrics, FM = Fidelity metrics).

NSS-based Metrics						
Metric's description	Symbol	DM	FM	Authors	Year	Ref.
Structural similarity based metrics						
Universal Quality Index	UQI		x	Wang et al.	2002	77
Structural Similarity Index	SSIM		x	Wang et al.	2003	78
Modified SSIM with automatic down-sampling	MSSIM		x	Wang et al.	2004	79
Multiscale SSIM	MS-SSIM		x	Rouse et al.	2004	80
Complex Wavelets-SSIM	CW-SSIM		x	Wang et al.	2005	81
Quality metric for Multichannel image based on SVD	M-SVD		x	Shnayderman et al.	2006	82
Right singular vector of the SVD-based measure	R-SVD		x	Mansouri et al.	2009	83
Local Standard Deviation Based Image Quality	LSDBIQ	x		Gore et al.	2015	86
Inf. theoretic based metrics						
Information Fidelity Criterion	IFC		x	Sheikh et al.	2005	84
Visual Information Fidelity index	VIF		x	Sheikh et al.	2006	85
Pixel-based VIF	P-VIF		x	Sheikh et al.	2006	85

and the correlation/monotonicity predictive performance parameters, respectively.

On the other side, the mid-level image features can be derived for a machine learning framework. Mid-level descriptors are interesting since they are typically close to image-level information with no attempt at high-level or structured image description. Examples of mid-level features which integrate some human visual factors include the edge amplitude, the edge length, the background activity and the background luminance. The very first IQA metric that exploited these descriptors has been proposed by Babu et al. in Ref. 108 where the authors used a radial basis function neural network. However, this metric is outside the scope of the present taxonomic scheme since it belongs to the NR category. Most recently, the Machine Learning-based Image Quality Measure (MLIQM) has been introduced by Charrier et al. to predict the visual quality of color images.[109] A vector of 25 image criteria is generated and trained in an SVM framework to predict the subjective scores provided in the image quality databases.

In addition to low-level features including contrast and structure factors calculated at five different levels, the luminance factor estimated on the achromatic component of the reference and the test images, a local chrominance distortion feature and a local colorimetric dispersion feature both calculated in the CIELab color space, a set of mid-level criteria has been derived in an attempt to modeling the HVS characteristics. Thereby, the steerable pyramid transform is performed with three levels and four orientation bands yielding to 12 low-pass filters.

Table IV. Summary of bio-inspired full-reference image quality metrics belonging to visual features-based methods from signal-driven category, where (DM = Distortion-aware metrics, FM = Fidelity metrics).

Visual Features-based Metrics						
Metric's description	Symbol	DM	FM	Authors	Year	Ref.
Spatial Domain Metrics						
Weighted Signal to Noise Ratio	WSNR		x	Mitsa et al.	1993	89
Picture Quality Scale	PQS	x		Miyahara et al.	1998	87
Noise Quality Metric	NQM	x		Damera-Venkata et al.	2000	88
Human Visual System-based Peak Signal to Noise Ratio	PSNR-HVS		x	Egiazarian et al.	2006	90
Visual Signal to Noise Ratio	VSNR		x	Chandler et al.	2007	15
Most Apparent Distortion	MAD		x	Larson et al.	2010	96
DCT-based metric combining texture masking and CSF	DCTex	x		Zhang et al.	2011	97
Transform Domain Metrics						
SNR using Wigner–Ville distribution	SNR_W		x	Iordache et al.	2001	92
SNR based on wavelet decomposition	SNR_{WAV}		x	Beghdadi et al.	2003	93
PSNR using Wigner–Ville distribution	$PSNR_W$		x	Beghdadi et al.	2006	94
Modified PSNR-HVS	PSNR-HVS-M		x	Ponomarenko et al.	2007	91
Riesz-transform-based feature similarity metric	RFSIM		x	Zhang et al.	2010	98
Feature similarity index	FSIM		x	Zhang et al.	2011	99

Machine learning oriented image quality methods have been extensively explored to develop RR and blind (NR) quality measures. They are known to be data-hungry, that is, they require a very large number of annotated image samples. The training stage of the design process is also known to be time-consuming.

DISCUSSION AND FUTURE TRENDS

This article has been dedicated to the description of a knowledge-based taxonomy for objective state-of-the-art FR image quality algorithms. A list of references interested in classification of IQMs has been supplied as well as a comprehensive listing of the literature for each topical area. The classification scheme is based on six kinds of knowledge of different levels: deep knowledge, intermediate knowledge

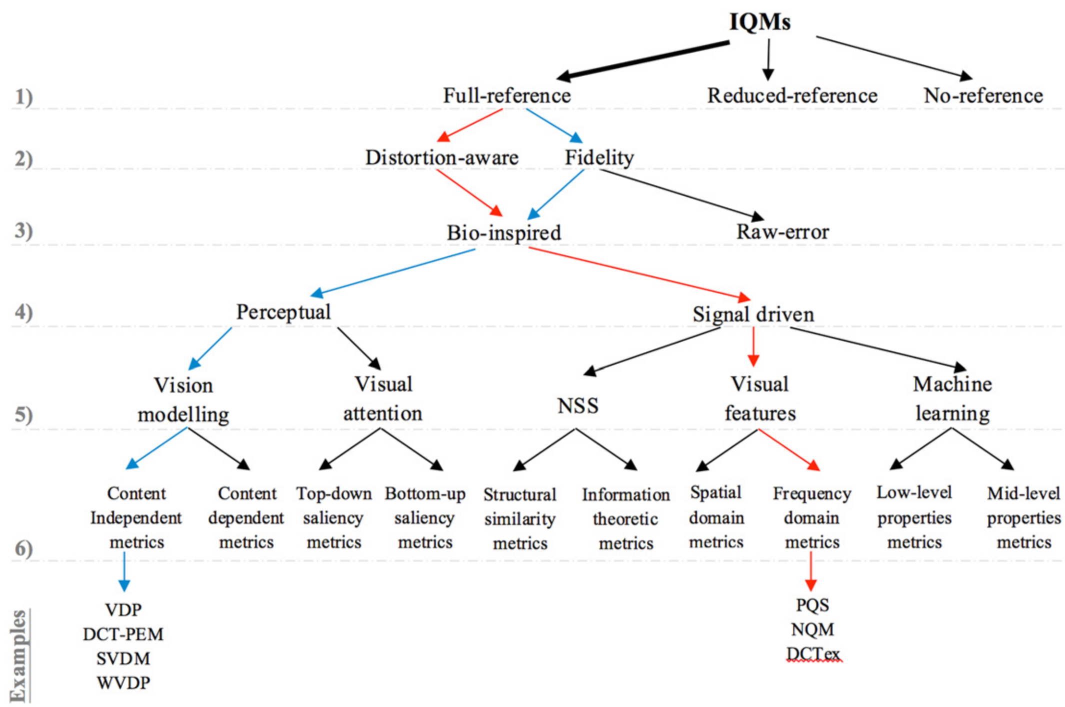

Figure 4. Examples of the use of the proposed knowledge-based taxonomic scheme.

Table V. Summary of bio-inspired full-reference image quality metrics belonging to machine learning oriented methods from signal-driven category, where (DM = Distortion-aware metrics, FM = Fidelity metrics).

Machine Learning Oriented Metrics						
Metric's description	Symbol	DM	FM	Authors	Year	Ref.
Low-level properties						
Circular back-propagation neural network-based metric	–		x	Carrai et al.	2002	100
MLP-based metric	–		x	Bouzerdoum et al.	2004	102
Circular back-propagation neural network-based metric	–		x	Gastaldo et al.	2005	101
SVM-based metric	–		x	Narwaria et al.	2010	103
SVM/MLP for feature selection-based metric	–		x	Lahoulou et al.	2010	106
Customized Cost Function-based image quality measure	ECF/CCF		x	Lahoulou et al.	2012	107
Mid-level properties						
Radial basis function neural network-based metric	—		x	Babu et al.	2007	108
Machine Learning-based Image Quality Measure	MLIQM		x	Charrier et al.	2012	109

and superficial level. The first layer of the proposed framework represents knowledge about the reference image. It allows outlining differences between FR IQMs on one side, and RR/NR metrics on the other side. The remaining layers are used to classify the FR image quality measures. They represent knowledge about degradations applied to images, visual perception field, HVS mechanisms, image analysis techniques, and finally knowledge about image representation and coding. The latter structure allows sorting out, in a more complete manner, the different theoretic foundations upon which the objective IQMs are built. Tables I–V are given with the aim to synthesize the evoked examples related to the proposed taxonomic scheme. They provide global and clear view on the fields that are generously explored and the ones where much work still to be accomplished. Figure 4 below represents a sketch on how to use the presented taxonomy and how to progress through the layers of knowledge. It is to note that the bold arrow corresponds to the path followed for this taxonomy for FR IQA metrics. For example, if we then follow the red arrows, we find the list of IQMs where (2) knowledge about the degradation is of higher level (distortion metrics), (3) knowledge about visual perception is deep (bio-inspired metrics), (4) knowledge about visual physiology and psychophysics is superficial (signal-driven approach), (5) knowledge about visual information analysis mechanisms is medium (visual features-based metrics), and (6) knowledge about perceptual image representation/coding is low level (transform domain class). One can make the same reasoning following the second example illustrated in Fig. 4 with blue arrows.

A natural extension of this work corresponds to the classification of RR and NR approaches as well as the classification of video quality assessment metrics.

The new trends in visual quality evaluation field follow three key directions, namely new applications, new content types and new models of human vision. Indeed, security related applications are of major concern in the almost all vision processing systems.[110] IQA is increasingly employed for liveness detection in biometric images. It allows the recognition of faked images used in attempt to violate the access right to secure physical or virtual areas. IQA has been extensively exploited for biometric, medical and satellite images. It has also been extended to 3D images and video and is about to be used for the complete and coherent summarization of visual data systems. Such schemes become absolutely necessary due to the huge amount of visual data available on the Internet which needs to be efficiently organized, exploited and analyzed. Another evolution concerns the understanding of the visual perception mechanisms. Several years ago, researchers were interested in the human vision as a phenomenon that occurs in the visual cortex ignoring thus the cerebral learning activity that often goes with the visual appreciation. Machine learning oriented category of IQA methods have been researched in order to include the experience in the perceptual quality judgment. However, the main shortcoming of those methods is that their predictive performance relies on the reliability of the subjective quality opinion scores available in the quality databases. Deep learning is being introduced in the field of visual quality assessment especially to design opinion-unaware metrics that have many implications for a broad range of applications. In addition to design models that can learn to predict human opinion without any prior knowledge, some endeavor is being directed toward distortion-unaware free-reference measures.[111] In fact, the NR algorithms were, for many years, degradation dependent; they rapidly became victims of early obsolescence due to the constant evolution of vision processing techniques.

REFERENCES

1 VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment — Phase II", *Rapport technique, Video Quality Experts Group* (2003)..

2 S. Winkler, "Analysis of public image and video databases for quality assessment," IEEE J. Sel. Top. Signal Process. **6**, 616–625 (2012) Special Issue on New Subjective and Objective Methodologies for Audio and Visual Signal Processing.

3 Z. Wang and A. C. Bovik, "Modern image quality assessment," *Syntheses Lectures on Image, Video and Multimedia Processing* (Morgan & Claypool Publishers, 2006).

4 D. M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," ISRN Signal Process. **2913**, 1–53 (2013).

5 J. Wang, G.- J. Qi, N. Sebe, and C. C. Aggarwal, "Guest editorial: big media data: understanding, search, and mining," IEEE Trans. Big Data **1**, 82–83 (2015).

6 Y. G. Jiang, "Categorizing Big Video Data on the Web: Challenges and Opportunities," *IEEE Int'l. Conf. on Multimedia Big Data (BigMM)* (IEEE, Piscataway, NJ, 2015), pp. 13–15.

7 M. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies," Proc. SPIE **5150**, 573–582 (2003).

8 S. Tourancheau, F. Autrusseau, Z. M. P. Sazzad, and Y. Horita, "Impact of subjective dataset on the performance of image quality metrics," *Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2008), pp. 365–368.

9 P. G. Engeldrum, "A short image quality model taxonomy," J. Imaging Sci. Technol. **48**, 160 (2004).

10 D. A. Silverstein and E. J. Farrell, "The relationship between image fidelity and image quality," *Proc. IEEE Int'l. Conf. on Image Processing* (IEEE, Piscataway, NJ, 1996), pp. 881–884.

11 I. Avcibas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," J. Electron. Imaging **11**, 206–223 (2002).

12 U. Engelke and H.-J. Zepernick, "Perceptual-based quality metrics for image and video services: a survey," *3rd European Conf. on Next Generation Internet Networks: Design Eng. Heterogeneity* (Trondheim, Norway, 2007), pp. 190–197.

13 X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment and human visual system," Proc. SPIE **7744**, 77440Z–17740Z (2010).

14 W. Lin and C.-C. Kuo, "Perceptual visual quality metrics: a survey," Elsevier Inc., J. Vis. Commun. Image Represent. **22**, 297–312 (2011).

15 D. M. Chandler and S. S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," IEEE Trans. Image Process. **16**, 2284–2298 (2007).

16 K.-H. Thung and P. Raveendran, "A survey of image quality measures," *Int'l. Conf. Technical Postgraduates TECHPOS* (Kuala Lumpur, 2009), pp. 14–15.

17 M. Pedersen and J. Y. Hardeberg, "Full-Reference image quality metrics: classification and evaluation," *Foundations and Trend in Computer Graphics and Vision* (2012), Vol. 7, pp. 1–80.

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

19 S. Akramullah, "Digital video concepts, methods, and metrics: quality, compression, performance, and power trade-off analysis," A. Press ed. (New York, 2014), pp. 101–160.

20 H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Process. **15**, 430–444 (2006).

21 K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," IEEE Trans. Image Process. **19**, 1427–1441 (2010).

22 K. Seshadrinathan and A. C. Bovik, "Automatic prediction of perceptual quality of multimedia signals-a survey," Int. J. Multimedia Tools and Applications, Special Issue on Survey Papers in Multimedia by World Experts **51**, 163–186 (2011).

23 A. K. Moorthy and A. C. Bovik, "Visual quality assessment algorithms: What does the future hold?," Int. J. Multimedia Tools and Applications, Special Issue on Survey Papers in Multimedia by World Experts **51**, 675–696 (2011).

24 A. Beghdadi, M.-C. Larabi, A. Bouzerdoum, and K. M. Iftekharuddin, "A survey of perceptual image processing methods," Signal Process., Image Commun. **28**, 811–831 (2013).

25 M.-C. Larabi, A. Saadane, and C. Charrier, "Quality assessment approaches," *Digital Color: Acquisition, Perception, Coding and Rendering* (Wiley Editions, France, 2012), pp. 265–306.

26 M.-C. Larabi, C. Charrier, and A. Saadane, "Quality assessment of still images," *Advanced Color Image Processing and Analysis* (Springer-Verlag, New York, 2013), pp. 423–447.

27 K. Seshadrinathan and A. C. Bovik, "The encyclopedia of multimedia," *Chapter Image and Video Quality Assessment* (Springer-Verlag, New York, 2009), pp. 8–17.

28 G. Ciocca, S. Corchs, F. Gasparini, and R. Schettini, "How to assess image quality within a workflow chain: an overview," Int. J. Digit. Libr. **15**, 1–25 (2014).

29 P. Mohammadi, A. Ebrahimi-Moghadam, and S. Shirani, "Subjective and objective quality assessment of image: a survey," Majlesi J. Electr. Eng. **9**, 55–83 (2015).

30 P. Le Callet and D. Barba, "A robust quality metric for color image quality assessment," *Int'l. Conf. on Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2003), Vol. 1, pp. 437–440.

31 H. R. Sheikh and A. C. Bovik, "An evaluation of recent full reference image quality assessment algorithms," IEEE Trans. Image Process. **15**, 3440–3451 (2006).

32 L. Zhang, L. Zhang, X. Mou, and D. Zhang, "A comprehensive evaluation of full reference image quality assessment algorithms," *19th IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2012), pp. 1477–1480.

33 A. Lahoulou, A. Bouridane, E. Viennet, and M. Haddadi, "Full-reference image quality metrics performance evaluation over image quality databases," Arab. J. Sci. Eng. **38**, 2327–2356 (2013).

34 M. Miyahara and K. Kotani, "Block distortion in orthogonal transform coding—analysis, minimization and distortion measure," IEEE Trans. Commun. **33**, 90–96 (1985).

35 S. A. Karunasekera and N. G. Kingsbury, "A distortion measure for blocking artifacts in images based on human visual sensitivity," IEEE Trans. Image Process. **4**, 713–724 (1995).
36 P. Fraënti, "Blockwise distortion measure for statistical and structural errors in digital images," Signal Process., Image Commun. **13**, 89–98 (1998).
37 A. Chalmers, A. McNamara, S. Daly, K. Myszkowski, and T. Troscianko, "Image quality metrics," *ACM SIGGRAPH Course Notes* (2000), p. 117.
38 Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? a new look at signal fidelity measures," IEEE Signal Process. Mag. **26**, 98–117 (2009).
39 Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," IET Electron. Lett. **44**, 800–801 (2008).
40 B. Girod, "What's wrong with mean squared error?," *Visual Factors of Electronic Image Communications*, edited by A. B. Watson (MIT Press, Cambridge, MA, 1993), pp. 207–220.
41 A. Ninassi, O. Le Meur, D. Barba, P. Le Callet, and A. Tirel, "Task impact on the visual attention in subjective image quality assessment," *European Signal Processing Conf. (EUSIPCO)* (Florence, Italy, 2006), pp. 1–5.
42 J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. Inf. Theory **20**, 525–536 (1974).
43 O. Faugeras, "Digital color image processing within the framework of a human vision model," IEEE Trans. Acoust. Speech Signal Process. **27**, 380–393 (1979).
44 F. W. Campbell and J. G. Robson, "Application of fourier analysis to the visibility of gratings," J. Physiol. **197**, 551–566 (1968).
45 A. B. Watson, "The cortex transform: rapid computation of simulated neural images," Comput. Vis. Graph. Image Process. **39**, 311–327 (1987).
46 S. Daly, "The visible difference predictor: an algorithm for the assessment of image fidelity," *Digital Images and Human Vision*, edited by A. B. Watson (The MIT Press, Cambridge, MA, 1993), pp. 179–206.
47 J. Lubin, "A visual discrimination model for image system design and evaluation," *Visual Models for Target Detection and Recognition*, edited by E. Peli (World Scientific Publishers, Singapore, 1995), pp. 207–220.
48 R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression," *IEEE Int'l. Conf. Acoust., Speech, and Signal Processing* (IEEE, Piscataway, NJ, 1989), pp. 1945–1948.
49 D. J. Heeger and P. C. Teo, "A model of perceptual image fidelity," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 1995), pp. 343–345.
50 S. Winkler, "A Perceptual Distortion Metric for Digital Color Images,," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 1998), pp. 399–403.
51 X. B. Gao, W. Lu, X. L. Li, and D. C. Tao, "Image quality assessment based on multiscale geometric analysis," IEEE Trans. Image Process. **18**, 1409–1423 (2009).
52 A. B. Watson, "DCT quantization matrices visually optimized for individual images," Proc. SPIE **1913**, 202–216 (1993).
53 A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," IEEE Trans. Image Process. **6**, 1164–1175 (1997).
54 W. Lu, X. B. Gao, D. C. Tao, and X. L. Li, "A wavelet-based image quality assessment method," Int. J. Wavelets, Multiresolution Inf. Process. **6**, 541–551 (2008).
55 A. P. Bradley, "A wavelet visible difference predictor," IEEE Trans. Image Process. **8**, 717–730 (1999).
56 Y. K. Lai and C.-C. J. Kuo, "A haar wavelet approach to compressed image quality measurement," J. Vis. Commun. Image Represent. **11**, 17–40 (2000).
57 A. B. Watson, "DCTune: a technique for visual optimization of DCT quantization matrices for individual images," *Society for Information Display Digest of Technical Papers XXIV* (1993), pp. 946–949.
58 M. Carnec, P. Le Callet, and D. Barba, "Objective quality assessment of color images based on a generic perceptual reduced reference," Signal Process., Image Commun. **23**, 239–256 (2008).
59 U. Engelke, H. Kaprykowsky, H. J. Zepernick, and P. Ndjiki-Nya, "Visual attention in quality assessment," IEEE Signal Process. Mag. **28**, 50–59 (2011).
60 W. Lin and L. Zhang, *Selective Visual Attention: Computational Models and Applications* (Wiley-IEEE Press, 2013).
61 A. Borji, D. N. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study," IEEE Trans. Image Process. **22**, 55–69 (2013).
62 H. Liu and I. Heynderickx, "Studying the added value of visual attention in objective image quality metrics based on eye movement data," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2009), pp. 3097–3100.
63 Q. Wang, L. Xu, Q. Chen, and Q. Sun, "Import of distortion on saliency applied to quality assessment," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2014), pp. 1165–1169.
64 W. Zhang, A. Borji, F. Yang, P. Jiang, and H. Liu, "Studying the added value of computational saliency in objective image quality assessment," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2014), pp. 21–24.
65 J. You, A. Perkis, M. Hannuksela, and M. Gabbouj, "Perceptual quality assessment based on visual attention analysis," *ACM Int'l. Conf. Multimedia* (Beijing, China, 2009), pp. 561–564.
66 E. C. Larson, C. Vu, and D. M. Chandler, "Can visual fixation patterns improve image fidelity assessment?," *IEEE Int'l. Conf. Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2008), pp. 2572–2575.
67 H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: based on eye-tracking data," IEEE Trans. Circuits Syst. Video Technol. **21**, 971–982 (2011).
68 W.-J. Tsai and Y.-S. Liu, "Foveation-based image quality assessment," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2014), pp. 25–28.
69 T. Yubing, H. Konik, F. A. Cheikh, and A. Tremeau, "Full reference image quality assessment based on saliency map analysis," J. Imaging Sci. Technol. **54**, 30503 (2010).
70 Q. Ma, L. Zhang, and B. Wang, "New strategy for image and video quality assessment," J. Electron. Imaging **19**, 11–19 (2010).
71 A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," IEEE J. Sel. Top. Signal Process. **3**, 193–201 (2009).
72 N. G. Sadaka, L. J. Karam, R. Ferzli, and G. P. Abousleman, "A no-reference perceptual image sharpness metric based on saliency-weighted foveal pooling," *15th IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2008), pp. 369–372.
73 L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1254–1259 (1998).
74 L. Zhang, Z. Gu, and H. Li, "SDSP: a novel saliency detection method by combining simple priors," *IEEE Int'l. Conf. Image Process. (ICIP)* (IEEE, Piscataway, NJ, 2013), pp. 171–175.
75 L. Zhang, Y. Shen, and H. Li, "VSI: a visual saliency-induced index for perceptual image quality assessment," IEEE Trans. Image Process. **23**, 4270–4281 (2014).
76 A. Saha and Q. M. Jonathan Wu, "Full reference image quality assessment by combining global and local distortion measures," IEEE Trans. Signal Process. **24**, 1–31 (2015).
77 Z. Wang and A. C. Bovik, "A universal image quality index," IEEE Signal Process. Lett. **9**, 81–84 (2002).
78 Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).
79 Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *37th IEEE Asilomar Conf. on Signals, Systems, and Computers* (IEEE, Piscataway, NJ, 2003), pp. 1398–1402.
80 D. M. Rouse and S. S. Hemami, "Analyzing the role of visual structure in the recognition of natural image content with multi-scale SSIM," Proc. SPIE **6806** (2008).
81 Z. Wang and E. P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," *IEEE Int'l. Conf. Acoustics, Speech and Signal Processing* (IEEE, Piscataway, NJ, 2005), pp. 573–576.
82 A. Shnayderman, A. Gusev, and A. M. Eskicioglu, "An SVD-based grayscale image quality measure for local and global assessment," IEEE Trans. image process. **15**, 422–429 (2006).
83 A. Mansouri, A. M. Aznaveh, F. Torkamani-Azar, and J. A. Jahanshahi, "Image quality assessment using the singular value decomposition theorem," Opt. Rev. **16**, 49–53 (2009).

[84] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," IEEE Trans. Image process. **14**, 2117–2128 (2005).
[85] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Process. **15**, 430–444 (2006).
[86] A. Gore and S. Gupta, "Full reference image quality metrics for JPEG compressed images," Int. J. Electron. Commun. (AEÜ) **69**, 604–608 (2015).
[87] M. Miyahara, K. Kotani, and V. R. Algazi, "Objective picture quality scale (PQS) for image coding," IEEE Trans. Commun. **46**, 1215–1225 (1998).
[88] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," IEEE Trans. Image Process. **9**, 636–650 (2000).
[89] T. Mitsa and K. L. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms," *IEEE Int'l. Conf. Acoustics, Speech and Signal Processing* (IEEE, Piscataway, NJ, 1993), pp. 301–304.
[90] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS," *2nd Int'l. Workshop on Video Processing and Quality Metrics* (Scottsdale, USA, 2006), p. 4.
[91] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," *3rd Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)* (Scottsdale, Arizona, USA, 2007), p. 4.
[92] R. Iordache and A. Beghdadi, "A Wigner-Ville distribution based image dissimilarity measure," *5th Int'l. Symposium on Signal Processing and its Applications (ISSPA)* (Kuala Lumpur, Malaysia, 2001), pp. 430–433.
[93] A. Beghdadi and B. Pesquet-Popescu, "A new image distortion measure based on wavelet decomposition," *6th Int'l. Symposium on Signal Processing and its Applications (ISSPA)* (Paris, France, 2003), pp. 485–488.
[94] A. Beghdadi and R. Iordache, "Image quality assessment using the joint space/spatial-frequency space," EURASIP J. Appl. Signal Process. **080537**, 1–8 (2006).
[95] D. Navon, "Forest before trees: the precedence of global features in visual perception," Cogn. Psychol. **9**, 353–383 (1977).
[96] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," J. Electron. Imaging **19**, 011006 (2010).
[97] F. Zhang, L. Ma, S. Li, and K. N. Ngan, "Practical image quality metric applied to image coding," IEEE Trans. Multimedia **13**, 615–624 (2011).
[98] L. Zhang, L. Zhang, and X. Mou, "RFSIM: a feature based image quality assessment metric using riesz transforms," *IEEE Int'l. Conf. Image Proc. (ICIP)* (IEEE, Piscataway, NJ, 2010), pp. 321–324.
[99] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," IEEE Trans. Image Process. **20**, 2378–2386 (2011).
[100] P. Carrai, I. Heynderickx, P. Gastaldo, and R. Zunino, "Image quality assessment by using neural networks," *IEEE Int'l. Symp. on Circuits & Systems, Sydney* (IEEE, Piscataway, NJ, 2002), pp. 253–256.
[101] P. Gastaldo, R. Zunino, I. Heynderickx, and E. Vicario, "Objective quality assessment of displayed images by using neural networks," Signal Process. Image Commun. **20**, 643–661 (2005).
[102] A. Bouzerdoum, A. Havstad, and A. Beghdadi, "Image quality assessment using a neural network approach," *4th IEEE Int'l. Symp. Signal Proc. Inf. Tech. (ISSPIT)* (IEEE, Piscataway, NJ, 2004), pp. 330–333.
[103] M. Narwaria and W. Lin, "Objective image quality assessment based on support vector regression," IEEE Trans. Neural Netw. **21**, 515–519 (2010).
[104] A. Lahoulou, E. Viennet, and A. Beghdadi, "Combining and selecting indicators for image quality assesment," *31st Int'l. Conf. Information Technology Interfaces* (Cavtat/Dubrovnik, Croatia, 2009), pp. 261–268.
[105] A. Lahoulou, E. Viennet, and A. Beghdadi, "Selecting low-level features for image quality assessment by statistical methods," J. Comput. Inf. Technol. CIT **18**, 83–195 (2010).
[106] A. Lahoulou, E. Viennet, and M. Haddadi, "Variable selection for image quality assessment using a neural network based approach," *2nd European Workshop on Visual Information Processing (EUVIP)* (Paris, France, 2010), pp. 45–50.
[107] A. Lahoulou, E. Viennet, and A. Bouridane, "Customizing cost function for optimizing image quality measures performances," Int. J. Electronics **99**, 1533–1546 (2012).
[108] R. V. Babu, S. Suresh, and A. Perkis, "No-reference JPEG image quality assessment using GAP-RBF," Signal Process. **87**, 1493–1503 (2007).
[109] C. Charrier, O. Lezoray, and G. Lebrun, "Machine Learning to design full-reference image quality assessment algorithm," Signal Process. Image Commun. **27**, 209–219 (2012).
[110] L. Janowski, M.-C. Leszczuk, M.-C. Larabi, and A. Ukhanova, "Optical recognition tasks," *Quality of Experience* (Springer International Publishing, 2014), pp. 383–394.
[111] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference approaches to image and video quality assessment," in *Multimedia Quality of Experience (QoE): Current Status and Future Requirements*, Ist ed. (John Wiley & Sons, Ltd., 2015), pp. 99–121, chap. 5.