https://doi.org/10.2352/ISSN.2169-2629.2017.32.7
Reprinted from
Journal of Imaging Science and Technology® 60(5): 050405-1–050405-13, 2016.

QUALITAS: Image Quality Assessment for Stereoscopic Images

Christine Fernandez-Maloigne▲
XLIM Laboratory, Signal, Image and Communications Department, University of Poitiers, 86962 Futuroscope, France

Jaime Moreno
XLIM Laboratory, Signal, Image and Communications Department, University of Poitiers, 86962 Futuroscope, France
ESIME, Instituto Politécnico Nacional, México

Alessandro Rizzi▲ **and Cristian Bonanomi**
Department of Computer Science, University of Milan, Via Comelico 39, 20135 Milan, Italy
E-mail: cristian.bonanomi@unimi.it

Abstract. *In this article, the authors present a method for assessing image quality in stereoscopic images: QUALITAS. The proposed method is inspired by some features of the human visual system, such as contrast sensitivity, response to visual disparity and perception of distance. Individual qualities of the stereo-pair are not simply averaged. QUALITAS introduces Contrast Band-Pass Filtering on a wavelet domain in both views; in this way it weights left and right images perceptually depending on viewing conditions. The authors have tested the method on the LIVE 3D stereoscopic image database and compared the results with a wide set of image quality metrics from current research.*

INTRODUCTION

Automatic or semiautomatic stereoscopic image quality assessment has arisen due to the recent diffusion of a new generation of stereoscopic technologies and content demand. However, no universally accepted method for stereoscopic image quality assessment exists. Considering all of the methods proposed in the literature, they can be mainly divided into three blocks, as presented in Figure 1: objective assessment, subjective assessment and strength of relationship.

In Fig. 1, we use 3D Coding in order to refer not only to Stereoscopic Coding but also to any coding that computes any 3D technology.

Objective assessment can be represented as a subsystem constituted by the following elements.

(1) Input: left and right images.
(2) Process: automatic assessment of stereoscopic coding features.
(3) Output: stereoscopic image + Stereoscopic Image Quality Assessment (SIQA).

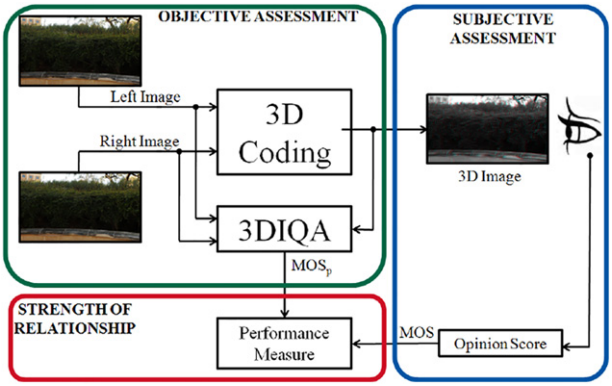


Figure 1. Block diagram of a general system for stereoscopic image quality assessment.

Here, the assessment verifies the efficacy of the process, in terms of quality. Therefore, the main goal of the SIQA is to measure in the stereoscopic image the quality and/or degradation of the original stereo-pair. In other words, any stereoscopic image coder could benefit from the support of SIQA in assessing its results. Therefore, it is important to highlight that the growth of SIQA algorithms is related to the intrinsic requirement to predict the quality in recent stereoscopic image coders.

Any SIQA algorithm is based on the evaluation of two or more views of the same scene, and the most of them use a 2D/Normal Image Quality Assessment (NIQA), in certain cases with the support of some model of Human Visual System (HVS) characteristics.

In some cases, basing SIQA on NIQA is straightforward. In many test scenarios this is not a problem, in particular where the test images utilized in psychophysical experiments or subjective assessments[1–3] have limited 3D volumes sometimes built as stages from a set of 2D scenarios.

SIQA algorithms can contribute to the prediction of not only a quality assessment, in general correlated with HVS characteristics, but also an estimate of the visual discomfort of the observer. Therefore, this article is intended not only for

▲ IS&T Members.

Received Apr. 8, 2016; accepted for publication June 30, 2016; published online Aug. 18, 2016. Associate Editor: Susan Farnand.

stereo image quality researchers but also for researchers who study the visual discomfort or classical NIQA algorithms.

Here, we propose a novel SIQA algorithm, called QUALITAS, which includes the main features of the HVS. Regarding its name, the QUALITAS metric involves the joint work of University of Poitiers (France), National Polytechnic Institute (Mexico) and University of Milan (Italy). The word "quality" in French, Spanish and Italian is *qualité*, *calidad* and *qualità*, respectively. These words are similar because they derive from the Latin word *qualitas*, hence the name.

This work is arranged in four sections. Qualitas Algorithm describes the work carried out in the field of subjective and objective stereoscopic image quality assessment. Then, in Experimental Results, the QUALITAS stereoscopic quality index is presented. Finally, in Conclusions, the experimental results are presented and discussed.

STEREOSCOPIC IMAGE QUALITY ASSESSMENTS

Objective Assessments

The main goal of any stereoscopic image quality assessment is to predict a subjective response (Fig. 1, green block, *Objective Assessment*).

Table I shows 25 stereoscopic metrics from 16 authors we consider in this article. This set of 25 metrics will be called SIQA-SET henceforth.

It is worth noting that some authors propose more than one metric, and we maintain their original metric name. Thus, we will refer to a certain metric by the name that appears in the corresponding row, not by its author. SIQA-SET was coded by ourselves in MATLAB.

We divide the SIQA-SET into two groups: the first group combines features of 2D/normal metrics, while the second group contains the metrics based on stereoscopic features. The first group of algorithms can use a generic 2D metric.

Therefore, the classification of SIQA-SET can be as follows.

Approaches Based on NIQA:

d_1, d_2, d_3, $\text{PSNR}_{\text{edge}}$, MSE_{dp}, YouDMOS_p and OQ.

Stereoscopic Approaches:

AkMOS_p, Ddl_1, Q_s, C_m, SBLC, ODDM_4, MSE_{ms}, PQM_{3D}, Q_{mao}, Q_{shao}, HDPSNR, 3VQM, IQA, SSA, DQ_{map1}, DQ_{map2}, DQ_{map3} and e_i.

The measures included in the SIQA-SET have been chosen based on their reported performance. In the same way we have collected 29 NIQAs in order to provide a baseline of normal metrics (NIQA-SET); then we can combine any of these 29 metrics with approaches of the SIQA-SET based on normal metrics, namely, d_1, d_2, d_3, $\text{PSNR}_{\text{edge}}$, MSE_{dp}, YouDMOS_p and OQ.

In the NIQA-SET we can find Statistical Image Quality Assessments (St-IQAs), Full-Reference Image Quality Assessments (FR-IQAs) and No-Reference Image Quality Assessments (NR-IQAs). The first 12 NIQAs are part of the MetrixMux toolbox,[20] while the rest of the metrics were collected from their respective authors.

Table I. Stereoscopic image quality assessments.

Algorithm	Metric
Akhter et al.[4]	AkMOS_p
Benoit et al.[5]	d_1
	d_2
	d_3
	Ddl_1
Bosc et al.[6]	Q_s
Chen et al.[7]	C_m
Gorlet et al.[8]	SBLC
Gu et al.[9]	ODDM_4
Hewage et al.[10]	$\text{PSNR}_{\text{edge}}$
Jin et al.[11]	MSE_{ms}
	MSE_{dp}
Joveluro et al.[12]	PQM_{3D}
Mao et al.[13]	Q_{mao}
Shao et al.[14]	Q_{shao}
Shen et al.[15]	HDPSNR
Solh et al.[16]	3VQM
Yang et al.[17]	IQA
	SSA
You et al.[18]	YouDMOS_p
	OQ
	DQ_{map1}
	DQ_{map2}
	DQ_{map3}
Zhu et al.[19]	e_i

St-IQA, FR-IQA and NR-IQA algorithms are listed for the sake of completeness, but not treated in this article, in order to contain its length. The reader can refer to the respective cited articles.

(1) Mean-Squared Error (MSE, St-IQA).
(2) Peak Signal-to-Noise Ratio (PSNR, St-IQA).
(3) Structural Similarity Index (SSIM, FR-IQA).[21]
(4) Multiscale SSIM Index (MSSIM, FR-IQA).[21]
(5) Visual Signal-to-Noise Ratio (VSNR, FR-IQA).[22]
(6) Visual Information Fidelity (VIF, FR-IQA).[23]
(7) Pixel-Based VIF (VIFP, FR-IQA).[24]
(8) Universal Quality Index (UQI, FR-IQA).[25]
(9) Image Fidelity Criterion (IFC, NR-IQA).[26]
(10) Noise Quality Measure (NQM, FR-IQA).[27]
(11) Weighted Signal-to-Noise Ratio (WSNR, FR-IQA).[28]
(12) Signal-to-Noise Ratio (SNR, St-IQA).
(13) Average Difference (AD, St-IQA).
(14) Maximum Difference (MD, St-IQA).
(15) Normalized Absolute Error (NAE, St-IQA).
(16) Normalized Cross Correlation (NCC, St-IQA).
(17) Structural Content (SC, St-IQA).
(18) Blind Image Quality Index (BIQI, NR-IQA).[29]

(19) Blind/Referenceless Image Spatial Quality Evaluator Index (BRISQUE, NR-IQA).[30]

(20) Naturalness Image Quality Evaluator (NIQE, NR-IQA).[31]

(21) No-Reference Peak Signal-to-Noise Ratio (NR-PSNR, NR-IQA).[32]

(22) Perceptual Peak Signal-to-Noise Ratio (P^2SNR, FR-IQA).[33]

(23) Feature-Similarity (FSIM, FR-IQA).[34]

(24) Riesz-Transform Feature-Similarity (RFSIM,FR-IQA).[35]

(25) Peak Signal-to-Noise Ratio with Contrast Sensitivity Function (PSNRHVSM, FR-IQA).[36]

(26) JPEG Quality Score (JQS, FR-IQA).[37]

(27) Practical Image Quality Metric (DCTEX, FR-IQA).[38]

(28) Most Apparent Distortion (MAD, FR-IQA).[39]

(29) Perceptual Quality Metric (PQM, FR-IQA).[12]

Subjective Assessments

In the field of subjective stereoscopic image quality assessment, few image databases have been developed. We have employed the LIVE 3D stereoscopic image database of the Laboratory for Image and Video Engineering of the University of Texas at Austin (USA) proposed by Moorthy et al.[1] LIVE 3D contains standardized psychophysical experiments,[40] and the stereoscopic image quality data are based on observer opinion score, collected with individual quality judgments (in Fig. 1 the blue block *Subjective Assessment*). In each trial, the images are rated on a scale of excellent, good, fair, poor and bad. Then, by means of statistical procedures, the data are processed, finally obtaining the mean opinion scores (MOSs). Each stereoscopic image database applies different statistical procedures; the reader can refer to the citation for the details. Additionally, the MOS merges results of different types in a form that allows the comparison with any stereoscopic assessment metric. Since SIQA predicts subjective responses, it obtains a predicted MOS or MOS_p.

Figure 2 shows the left views of 20 reference images used in this subjective assessment. On the other hand, we want to mention its main features. Table II depicts the main characteristics of LIVE 3D. The distortions for LIVE 3D are JPEG2000 (JP2K), JPEG, Additive White Gaussian Noise (WN), Gaussian Blur (Blur) and Fast-Fading (FF).

It is worth mentioning that this database is not the only work carried out in this field. Goldmann et al.,[2] Wang et al.,[3] Park et al.[41] and Huan et al.[42] proposed other databases intended to improve the stereoscopic image quality. All four image databases have different sizes of reference and distorted images. Thus, resizing of the images could change the precision of the subjective results.

This article compares the psychophysical experiments of the LIVE 3D image database against a collected set of SIQAs.

Table II. Principal Features of the LIVE 3D stereoscopic image database.

Reference images	20
Distorted images	365
Image format	bitmap (BMP)
Studio images	No
Resolution	640 × 360
Views	2
Distortions	5
Observers	32
Camera	1 × Nikon D700
Capture process	Single shot
3D display	Viewsonic IZ3D

QUALITAS ALGORITHM

Theoretical Framework

Contrast Band-Pass Filtering

Contrast Band-Pass Filtering (CBPF) is usually used in digital image processing to implement the contrast sensitivity function of the HVS in a simplified way. It estimates the image perceived by an observer at a distance d just by modeling the chromatic sensitivity of our vision. That is, given an image I and an observation distance d, CBPF obtains an estimation of what is perceived when observing I at distance d. CBPF is based on three stimulus properties: spatial frequency, spatial orientation and surround contrast.

The sensitivity of a human observer to contrast with respect to spatial frequency is described by the Contrast Sensitivity Function (CSF). CBPF is usually used in digital image processing to implement this sensibility to spatial contrast in a simplified way. In our model, we want to take into account the strength of image regions within wavelet decomposition. Therefore, the accuracy of the different regions, within the decomposition in sub-bands, which occupy an interval on the spatial frequencies, is adjusted according to the contrast sensitivity. A further advantage of this procedure is the possibility to model the different responses of the HVS, according to the considered frequency band.

The CBPF model takes an input image I and decomposes it into a set of wavelet planes $\omega_{s,o}$ of different spatial scales s and spatial orientations o. It is described as

$$I = \sum_{s=1}^{n} \sum_{o=v,h,\mathrm{dgl}} \omega_{s,o} + c_n, \tag{1}$$

where n is the number of wavelet planes, c_n is the residual plane and o is the spatial orientation, namely, vertical (v), horizontal (h) or diagonal (dgl).

The filtered image I_ρ is recovered by weighting these $\omega_{s,o}$ wavelet coefficients. The weighting function considers spatial surround information (denoted by r), visual frequency (ν) and observation distance (d). The filtered image I_ρ is

Figure 2. Left views of the 20 reference images used in the subjective study (LIVE 3D) of Moorthy et al.[1]

obtained by

$$I_\rho = \sum_{s=1}^{n} \sum_{o=v,h,\mathrm{dgl}} \alpha\,(v,r)\,\omega_{s,o} + c_n, \tag{2}$$

where $\alpha(v,r)$ is the weighting function which aims to reproduce the HVS properties discussed above. Here, $\alpha(v,r)\omega_{s,o} \equiv \omega_{s,o;\rho,d}$ can be considered to be the *perceptual wavelet coefficient* of image I when observed at distance d and is written as

$$\alpha(v,r) = z_{\mathrm{ctr}} C_d(\hat{s}) + C_{\min}(\hat{s}), \tag{3}$$

where the three terms that describe $\alpha(v,r)$ are defined as follows.

z_{ctr}: non-linear function and estimation of the local difference relative to its surround, ranging from zero to one, defined by

$$z_{\mathrm{ctr}} = \frac{\left[\dfrac{\sigma_{\mathrm{cen}}}{\sigma_{\mathrm{sur}}}\right]^2}{1+\left[\dfrac{\sigma_{\mathrm{cen}}}{\sigma_{\mathrm{sur}}}\right]^2}, \tag{4}$$

where σ_{cen} and σ_{sur} are the standard deviations of the wavelet coefficients in two concentric rings, which represent a center–surround interaction around each coefficient.

$C_d(\hat{s})$: weighting function defined as a piecewise Gaussian function, such as

$$C_d(\hat{s}) = \begin{cases} e^{-\hat{s}^2/2\sigma_1^2}, & \hat{s} = s - s_{\mathrm{thr}} \le 0, \\ e^{-\hat{s}^2/2\sigma_2^2}, & \hat{s} = s - s_{\mathrm{thr}} > 0. \end{cases} \tag{5}$$

$C_{\min}(\hat{s})$: term that avoids the $\alpha(v,r)$ function being zero and that is defined by

$$C_{\min}(\hat{s}) = \begin{cases} \frac{1}{2} e^{-\hat{s}^2/2\sigma_1^2}, & \hat{s} = s - s_{\mathrm{thr}} \le 0, \\ \frac{1}{2}, & \hat{s} = s - s_{\mathrm{thr}} > 0, \end{cases} \tag{6}$$

taking $\sigma_1 = 2$ and $\sigma_2 = 2\sigma_1$. Both $C_{\min}(\hat{s})$ and $C_d(\hat{s})$ depend on the factor s_{thr}. The latter factor is the scale associated with 4 cpd when an image is observed from the distance d; l_p is the pixel size at one visual degree,[43] defined by

$$s_{\mathrm{thr}} = \log_2\left(\frac{d\tan(1^\circ)}{4l_p}\right). \tag{7}$$

Figure 3 shows two examples of CBPF images of im2_l (left view of second image from the LIVE 3D image database), calculated by Eq. (2) for a 19 inch monitor with 1280 pixels of horizontal resolution, at $d = \{100, 1000\}$ centimeters. CBPF is a piecewise Gaussian function; therefore, at long distances, such as 1000 cm, the filtered image seems blurred, Fig. 3(b).

Quality Assessment

The employed image quality assessment is inspired by the algorithm of Wang et al.[25] and is described by

$$\mathrm{QA}[f(i,j), \hat{f}(i,j)] = \mathrm{LC} + \mathrm{LumD} + \mathrm{CnD}, \tag{8}$$

where $f(i,j)$ represents the original reference image and $\hat{f}(i,j)$ represents a distorted version of $f(i,j)$ (whose quality in comparison to $f(i,j)$ is being evaluated).

Thus, QA is an objective quality metric, which is based on the interaction of three terms: Loss of Correlation (LC), Luminance Distortion (LumD) and Contrast Distortion (CnD).

(a) (b)

Figure 3. Filtered images of im2_l obtained by CBPF at different observation distances: (a) $d = 100$ cm, (b) $d = 1000$ cm.

(a) (b) (c) (d)

Figure 4. (a), (b) LIVE 3D stereo-pair 5 with JPEG distortion level = 3, $R = 0.9909$. (c), (d) LIVE 3D stereo-pair 17 with JPEG distortion level = 4 (c) and distortion level = 1 (d), $R = 0.9655$.

These terms are described as follows:

$$\begin{aligned} \mathrm{LC} &= \frac{\mathrm{cov}[f(i,j), \hat{f}(i,j)]}{\mathrm{var}[f(i,j)] \times \mathrm{var}[\hat{f}(i,j)]}, \\ \mathrm{LumD} &= \frac{2 \times \mathrm{mean}[f(i,j)] \times \mathrm{mean}[\hat{f}(i,j)]}{\mathrm{mean}[f(i,j)]^2 + \mathrm{mean}[\hat{f}(i,j)]^2}, \\ \mathrm{CnD} &= \frac{2 \times \mathrm{std}[f(i,j)] \times \mathrm{std}[\hat{f}(i,j)]}{\mathrm{var}[f(i,j)] + \mathrm{var}[\hat{f}(i,j)]}, \end{aligned} \tag{9}$$

where cov is the sample covariance between $f(i,j)$ and $\hat{f}(i,j)$, while var and mean are, respectively, the sample variance and the sample mean of either $f(i,j)$ or $\hat{f}(i,j)$.

Weighting Function Based on Wavelet Energy Ratio

In this article, we also propose a weighing function intended to balance effects of unequal bit-allocation to left and right views based on wavelet energy ratio. These kinds of effects have been described by several authors, such as Stelmach et al.,[44] Palaniappan et al.[45] Vatolin et al.[46] and Chen et al.[47]

The total energy measure or the *deviation signature*[48] ε is the absolute sum of the wavelet coefficient magnitudes, defined by Wilson et al.[49] as

$$\varepsilon = \sum_{n=1}^{N} \sum_{m=1}^{M} |x(m,n)|, \tag{10}$$

where $x(m,n)$ is the set of wavelet coefficients, whose energy is being calculated.

Therefore, Eq. (11) expresses the relative wavelet energy ratio R, which compares how different the energy of the left image is compared with the right one. This ratio R also introduces features of visual discomfort to the QUALITAS approach:

$$R = \frac{\varepsilon|f_l(i,j)|}{\varepsilon|f_r(i,j)|}, \tag{11}$$

where $f_l(i,j)$ and $f_r(i,j)$ are the original stereo-pair. When the energy of the left image is lower than or equal to that of the right view, the dynamic range of R is [0,1]. Otherwise, when the energy of the right image is lower than that of the left view, Eq. (11) has to be inverted in order to maintain the dynamic range, then $R = R^{-1}$.

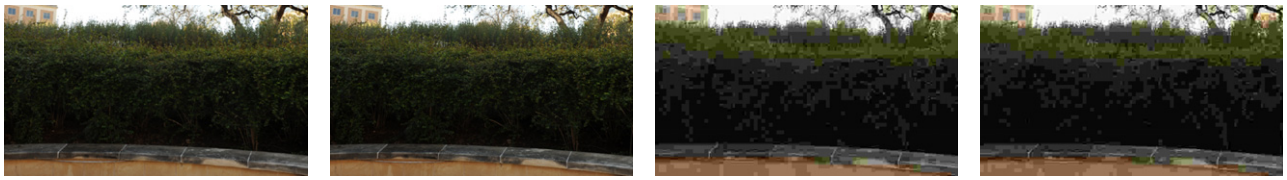
Figures 4(a) and (b) show LIVE 3D stereo-pair no. 5 with JPEG distortion level = 3 to both the left and right views (balanced distortion), with $R = 0.9909$, while Figs. 4(c) and (d) depict LIVE 3D stereo-pair no. 17 with JPEG distortion level = 4 for the left view and distortion level = 1 for the right view (unbalanced distortion), respectively, with $R = 0.9655$. Thus, we can conclude that when the level of distortion is similar, regardless of the degree of distortion, R would tend to be 1. Otherwise, if the level of distortion is different, R would tend to be 0.

Stereoscopic Quality Assessment

Our proposed stereoscopic image quality assessment is defined as follows:

$$\mathrm{QUALITAS} = \left[\sum_{i=l,r} \sum_{p=F,B} \frac{\mathrm{QA}(\mathrm{FOr}_{i,p}, \mathrm{FDis}_{i,p})}{4} \right]^{R}, \tag{12}$$

where i is either the l or the r image, i.e., the left or right view, respectively. The term p is related to the plane, either

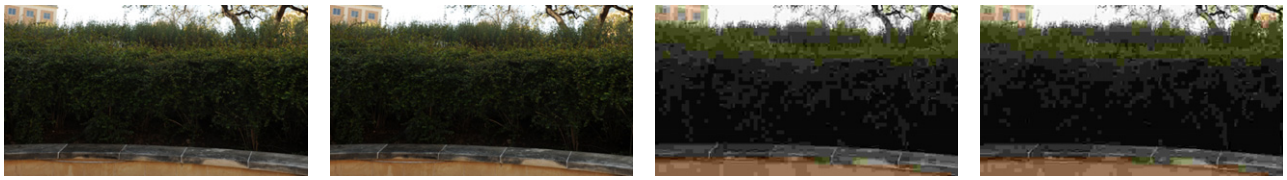


(a) (b) (c) (d)

Figure 5. (a), (b) Original stereo-pair; (c), (d) distorted stereo-pair.

(a) (b) (c)

Figure 6. (a) Disparity map dM. (b), (c) Binary masks Msk_F for foreground (b) and Msk_B for background (c).



(a) (b) (c) (d)

Figure 7. (a), (b) Filtered original stereo-pair (${}_{i=\lambda p}\text{FOr}_i$). (c), (d) Filtered distorted stereo-pair (${}_{-i\lambda p}\text{FDis}_i$).

foreground (F) or background (B), of the stereoscopic scene. The terms $\text{FOr}_{i,p}$ and $\text{FDis}_{i,p}$ are the original and distorted contrast pass-band filtered stereo-pairs, respectively, of the ith image in the pth plane, while R is the energy ratio in the original stereo-pair.

Thus, QUALITAS can be considered as a methodology divided into five steps. In order to exemplify our methodology, we consider $f_l(i,j)$ (Figure 5(a)) and $f_r(i,j)$ (Fig. 5(b)) as the original stereo-pair and $\hat{f}_l(i,j)$ (Fig. 5(c)) and $\hat{f}_r(i,j)$ (Fig. 5(d)) as the distorted stereo-pair.

Dividing the Stereoscopic Scene.
Let $f_l(i,j)$ and $f_r(i,j)$ represent the original stereo-pair to be compared against a distorted version of it, being the left and right views, respectively.

Then, the disparity map dM is computed using $f_l(i,j)$ and $f_r(i,j)$ (Figure 6(a)). It is noteworthy that the LIVE 3D image database release contains both left and right display maps, and one or both of them can be used. Moreover, dM represents the apparent distances in gray scale, so $dM(\boldsymbol{d})$ is a matrix of these distances in centimeters. In order to bisect the stereoscopic scene into foreground and background planes, the mean distance $\bar{d}$ in $dM(\boldsymbol{d})$ is calculated as

$$\bar{d} = \frac{[dM(d)]_{\max} - [dM(d)]_{\min}}{2}. \tag{13}$$

Thus, if a given distance in $dM(\boldsymbol{d})$ is less than $\bar{d}$, then it is considered as a part of the foreground mask (Msk_F, Fig. 6(b)), otherwise it is considered as a part of the background mask (Msk_B, Fig. 6(c)).

Contrast Band-Pass Filtering.
Let us filter the distorted and original stereo-pairs by means of Eq. (2); thus FOr_l, FOr_r, FDis_l and FDis_r are filtered images, which maintain the high frequencies of $f_l(i,j)$, $f_r(i,j)$, $\hat{f}_l(i,j)$ and $\hat{f}_r(i,j)$, respectively, at a distance of $\bar{d}$ centimeters from the observer.

Figures 7(a) and (b) show the filtered original stereo-pair ($\sum_{i=l,r}\text{FOr}_i$), while Fig. 7(c) and (d) show the filtered distorted stereo-pair ($\sum_{i=l,r}\text{FDis}_i$).

Depth Segmentation.
We apply masks Msk_F and Msk_B to both $\sum_{i=l,r}\text{FOr}_i$ and $\sum_{i=l,r}\text{FDis}_i$, obtaining eight segmented images, shown in Figure 8. Thus, the filtered and segmented original stereo-pair can be defined as $\sum_{i=l,r}\sum_{p=F,B}(\text{FOr}_{i,p})$ and the filtered and segmented distorted stereo-pair as $\sum_{i=l,r}\sum_{p=F,B}(\text{FDis}_{i,p})$.

Quality Assessment.
We evaluate separately the quality of distorted filtered and segmented stereo-pairs employing Eq. (8), i.e.,

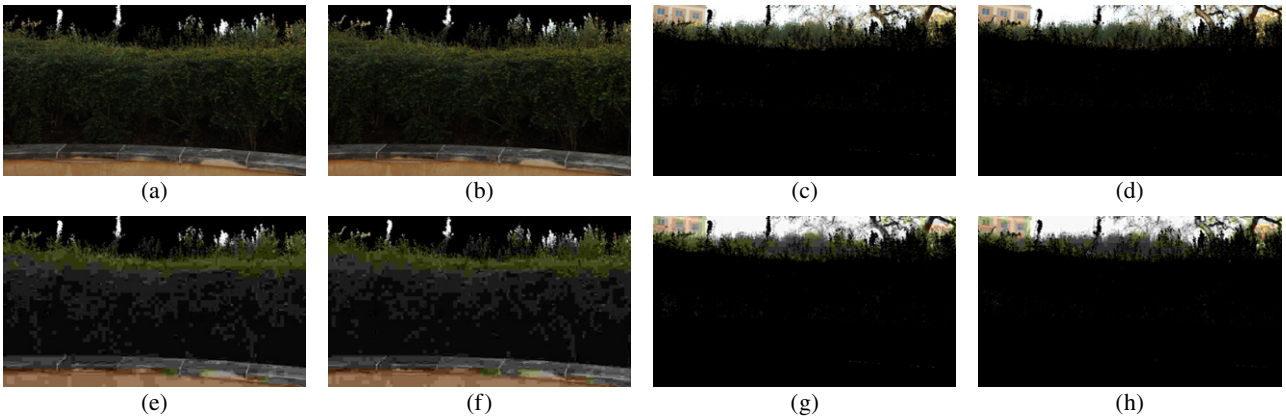


Figure 8. (a)–(d) Filtered and segmented original stereo-pair ($\sum_{i=l,r}\sum_{p=F,B}(\mathrm{FOr}_{i,p})$). (e)–(h) Filtered and segmented distorted stereo-pair ($\sum_{i=l,r}\sum_{p=F,B}(\mathrm{FDis}_{i,p})$).

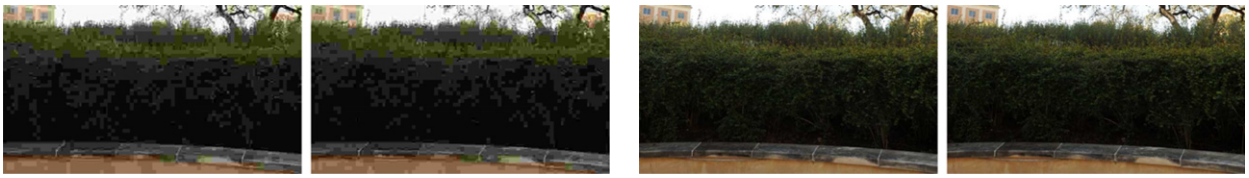


Figure 9. (a) Highly distorted LIVE 3D stereo-pair 2 with JPEG distortion, QUALITAS = 0.6814. (b) Slightly distorted LIVE 3D stereo-pair 2 with JPEG distortion, QUALITAS=0.9322.

$\sum_{i=l,r}\sum_{p=F,B} QA(\mathrm{FOr}_{i,p}, \mathrm{FDis}_{i,p})$. Then, we weight the obtained scores by 1/4 in order to adjust the dynamic range to [0,1].

Effect of Disparate Image Quality.
Finally, Eq. (11) is applied to $\sum_{i=l,r}\sum_{p=F,B} QA(\mathrm{FOr}_{i,p}, \mathrm{FDis}_{i,p})$ in order to balance any possible effect of unequal bit-allocation into the left and right original stereo-pairs.

Thus, in this example, QUALITAS predicts for Figure 9(b) a quality score of 0.9322, which corresponds to *Low JPEG Distortion*, and for Fig. 9(a) a score of 0.6814, which corresponds to *High JPEG Distortion*.

Relation of QUALITAS with Recent Works

QUALITAS is a set of algorithms which predict the quality of a stereo-pair. It includes different aspects from 3D quality metrics validated in the literature. We can mention these different aspects in relation with other recent works.

(1) Dividing the stereoscopic scene: Yasakethu et al.[50,51] segment into three planes to predict the sensation of depth in 3D video, while we divide the image into two planes, background and foreground.
(2) Contrast Band-Pass Filtering: Otazu et al.[43] propose an Extended Contrast Sensitivity Function. We modify this function to filter the two planes according to the distance of the observer.
(3) Depth segmentation: Bosc et al.,[6] Shen et al.[15] and Yang et al.[17] employ different kinds of depth segmentations. We adapt these algorithms for the two filtered planes of the stereo-pair.
(4) Quality assessment: Wang et al.[25] propose a quality index for the 2D/normal image. We have adapted this approach to segment a localized set of pixels.
(5) Effect of disparate image quality: Moreno[33] measures the amount of perceptual energy of a 2D/normal perceptual image quality assessment. We have adapted this function to measure the ratio of the difference of the energy in the original stereo-pair.

EXPERIMENTAL RESULTS

The evaluation results of every observer group (*MOS*) and image quality metric (MOS_p) are normalized using the following equation:

$$\hat{MOS}_p = \frac{MOS_p - MOS_p^{\min}}{MOS_p^{\max} - MOS_p^{\min}}, \tag{14}$$

where MOS_p denotes the calculated value of each metric, and $MOS_p^{\min}$ and $MOS_p^{\max}$ are the minimum and maximum values predicted in the whole LIVE 3D image database. We also employ Eq. (14) for normalizing *MOS* results of the LIVE 3D image database.

From Fig. 1, red block, the strength of the relationship between the normalized *MOS* and MOS_p is measured by a Performance Measure (PM), as a correlation coefficient between the two results. Strength of Relationship indicates the tendency of these two metrics to move in the same or

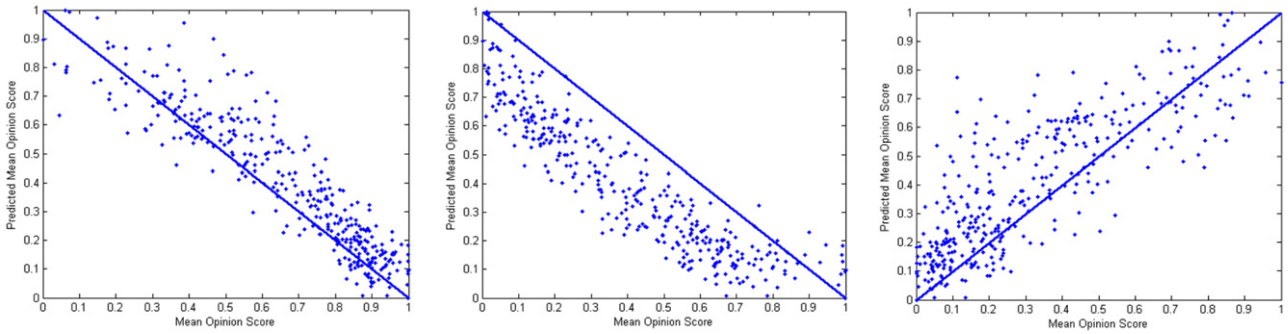


Figure 10. MOS versus MOS_p (both normalized). MOS_p is predicted by (a) d_1 using FSIM, (b) d_2 using UQI and (c) DQ_{map2}.

opposite direction. The performance measures used are as follows:

- Spearman's Rank Ordered Correlation Coefficient (SROCC),
- Kendall's Rank Ordered Correlation Coefficient (KROCC),
- Pearson's Linear Correlation Coefficient (LCC),
- Root-Mean-Squared Error (RMSE).

Any correlation coefficient value close to 1 indicates good correlation between the two. In addition, RMSE indicate good performance. The Pearson correlation is $+1$ in the case of a perfect increasing linear relationship and -1 in the case of a perfect decreasing one (anticorrelation). When it approaches zero, it means an uncorrelation. The closer the coefficient is to either -1 or $+1$, the stronger the correlation is between the variables. The closer the coefficient is to 0 the more independent the variables are. RMSE is a complementary measure that associated with the previous three completes the score of the test.

Furthermore, we present our results in the following three ways.

- Scatter plots depict the relationship between subjective results (normalized MOS) and objective results (normalized MOS_p) of a certain SIQA, in Figures 10 and 11.
- Tables III–V show the overall results of the strength of relationship of a set of SIQAs computed on the LIVE 3D image database, together with five different distortions:
  (1) ringing artifacts of a JPEG2000k image compression (JP2K),
  (2) blocking artifacts of a JPEG image compression (JPEG),
  (3) additive white Gaussian noise (WN),
  (4) Gaussian blurring (Blur),
  (5) fast fading noise (FF).
- Tables VI–IX show separately the best performance results of the whole SIQA-SET compared with QUALITAS.

Table III. Overall performance across SIQA-SET in predicting perceived stereoscopic image quality: Linear Correlation Coefficient (LCC), Spearman's Rank Ordered Correlation Coefficient (SROCC), Kendall's Rank Ordered Correlation Coefficient (KROCC), Root Mean Squared Error (RMSE).

Distortion	SIQA	NIQA	PM	Value
ALL	d_1	FSIMC	LCC	0.9169
	d_2	UQI	SROCC	0.9335
	d_2	UQI	KROCC	0.7659
	DQ_{map2}	none	RMSE	0.1289
JP2K	d_2	UQI	LCC	0.9304
	d_2	UQI	SROCC	0.9104
	d_2	UQI	KROCC	0.7405
	DQ_{map2}	none	RMSE	0.0961
JPEG	d_2	UQI	LCC	0.7620
	d_2	UQI	SROCC	0.7268
	d_2	UQI	KROCC	0.5212
	DQ_{map2}	none	RMSE	0.0742
WN	Ddl_1	none	LCC	0.9330
	d_2	MSSIM	SROCC	0.9425
	d_2	MSSIM	KROCC	0.7911
	d_1	BRISQUE	RMSE	0.1001
Blur	d_2	UQI	LCC	0.9558
	MSE_{ms}	none	SROCC	0.9318
	$YouDMOS_p$	AD	KROCC	0.7818
	$PSNR_{edge}$	NAE	RMSE	0.1156
FF	d_2	UQI	LCC	0.8549
	d_2	UQI	SROCC	0.8162
	d_2	UQI	KROCC	0.6245
	d_2	BPSNR	RMSE	0.1116

SIQA-SET

Table III shows the performance of an overall test, which includes all SIQAs of the SIQA-SET described in Objective Assessments. SIQA-SET is a compound set of 221 metrics, since we consider 203 metrics combining d_1, d_2, d_3, $PSNR_{edge}$, MSE_{dp}, $YouDMOS_p$ and OQ with NIQA algorithms (seven 3D metrics combined with 29 2D metrics), in addition to 18 purely stereoscopic approaches.

Table IV. Overall performance of QUALITAS in predicting perceived stereoscopic image quality: Linear Correlation Coefficient (LCC), Spearman's Rank Ordered Correlation Coefficient (SROCC), Kendall's Rank Ordered Correlation Coefficient (KROCC), Root Mean Squared Error (RMSE).

Distortion	SIQA	PM	Value
ALL	QUALITAS	LCC	0.9392
	QUALITAS	SROCC	0.9334
	QUALITAS	KROCC	0.7668
	QUALITAS	RMSE	0.4754
JP2K	QUALITAS	LCC	0.9467
	QUALITAS	SROCC	0.9126
	QUALITAS	KROCC	0.7443
	QUALITAS	RMSE	0.4907
JPEG	QUALITAS	LCC	0.7557
	QUALITAS	SROCC	0.7384
	QUALITAS	KROCC	0.5396
	QUALITAS	RMSE	0.5838
WN	QUALITAS	LCC	0.9333
	QUALITAS	SROCC	0.9336
	QUALITAS	KROCC	0.7703
	QUALITAS	RMSE	0.4539
Blur	QUALITAS	LCC	0.9397
	QUALITAS	SROCC	0.9290
	QUALITAS	KROCC	0.7697
	QUALITAS	RMSE	0.4799
FF	QUALITAS	LCC	0.8684
	QUALITAS	SROCC	0.8256
	QUALITAS	KROCC	0.6371
	QUALITAS	RMSE	0.3705

Table V. Overall performance across SIQA-SET including QUALITAS in predicting perceived stereoscopic image quality: Linear Correlation Coefficient (LCC), Spearman's Rank Ordered Correlation Coefficient (SROCC), Kendall's Rank Ordered Correlation Coefficient (KROCC), Root Mean Squared Error (RMSE).

Distortion	SIQA	NIQA	PM	Value
ALL	QUALITAS	none	LCC	0.9392
	d_2	UQI	SROCC	0.9335
	QUALITAS	none	KROCC	0.7668
	DQ_{map2}	none	RMSE	0.1289
JP2K	QUALITAS	none	LCC	0.9467
	QUALITAS	none	SROCC	0.9126
	QUALITAS	none	KROCC	0.7443
	DQ_{map2}	none	RMSE	0.0961
JPEG	d_2	UQI	LCC	0.7620
	QUALITAS	none	SROCC	0.7384
	QUALITAS	none	KROCC	0.5396
	DQ_{map2}	none	RMSE	0.0742
WN	QUALITAS	none	LCC	0.9333
	d_2	MSSIM	SROCC	0.9425
	d_2	MSSIM	KROCC	0.7911
	d_1	BRISQUE	RMSE	0.1001
Blur	d_2	UQI	LCC	0.9558
	MSE_{ms}	none	SROCC	0.9318
	$YouDMOS_p$	AD	KROCC	0.7818
	$PSNR_{edge}$	NAE	RMSE	0.1156
FF	QUALITAS	none	LCC	0.8684
	QUALITAS	none	SROCC	0.8256
	QUALITAS	none	KROCC	0.6371
	d_2	BPSNR	RMSE	0.1116

Table III shows the best results in terms of linear and non-linear correlation of these 221 metrics and the rest of the SIQAs not based on NIQA, with all 365 images of the LIVE 3D image database. We use ← to indicate that we applied a certain NIQA algorithm to a SIQA. Then, the results show that the best linear correlation is obtained by $d_1 \leftarrow$ FSIM (91.69%); see also Fig. 10(a). Meanwhile, $d_2 \leftarrow$ UQI is the best ranking metric, since it obtains the best correlation with human observers regarding both SROCC and KROCC. Moreover, based on the results of DQ_{map2}, it is clear that for the set of distortions considered, this metric is the most accurate (Fig. 10(c)). However, considering only image compression distortions, JPEG2000k and JPEG, we can highlight that $d_2 \leftarrow$ UQI is the best metric in either linear or rank correlations.

QUALITAS

Table IV shows the QUALITAS results on distortions both taken one by one and in combination. These results show that QUALITAS is highly correlated with observer results. When all images of LIVE 3D are compared, LCC, SROCC and KROCC are, respectively, 93.92%, 93.34% and 76.68%, and RMSE= 0.4754. Addition of a factor that weighs the relative wavelet energy ratio R increases the degree of linear correlation. Thus, incorporation of R yields good results not only on the average response of all distortions, but also on individual distortions such as JPEG2000 or additive white Gaussian noise.

Fig. 11 shows the points highly concentrated around the diagonal line of maximal linear correlation.

SIQA-SET versus Qualitas

Table V shows the performance of an overall experimental test, which includes all SIQAs of the SIQA-SET and QUALITAS.

It can be noticed from Fig. 11 that QUALITAS correlates with MOS at 93.92% in terms of LCC and 76.68% in terms of KROCC, scoring the best rank among the SIQA-SET. In terms of SROCC, the best-ranked metric is $d_2 \leftarrow UQI$ with 93.35. Finally, for JPEG2000 and JPEG distortions, QUALITAS also obtains the highest rank considering both compression distortions in both rank ordered coefficients.

Therefore, we have presented several metrics (221) to evaluate the quality of a stereo-pair. Researchers interested in stereoscopic coding, visual discomfort or stereoscopic display would be interested in finding the best metric taking

Table VI. Performance across SIQA-SET in predicting perceived stereoscopic image quality: Linear Correlation Coefficient (LCC). Bold indicates the best metric, while italics indicate the second best.

Author	SIQA	NIQA	JP2K	JPEG	WN	Blur	FF	ALL
Benoit et al.[5]	d_1	FSIM	*0.9119*	*0.6259*	*0.9307*	*0.9358*	0.7834	*0.9169*
Bosc et al.[6]	Q_s	none	0.0259	0.1563	0.8866	0.1853	0.0882	0.4115
Gu et al.[9]	ODDM_4	none	0.7728	0.4461	0.9223	0.7024	0.7540	0.7460
Hewage et al.[10]	$\text{PSNR}_{\text{edge}}$	NCC	0.6737	0.3293	0.7997	0.8027	0.7738	0.5772
Jin et al.[11]	MSE_{dp}	UQI	0.8512	0.5769	0.8832	0.8523	0.6327	0.7962
Joveluro et al.[12]	PQM_{3D}	none	0.1393	0.2415	0.8477	0.0444	0.1765	0.4790
Mao et al. [13]	Q_{mao}	none	0.7189	0.1290	0.7701	0.7527	0.4413	0.7082
Yang et al.[17]	*IQA*	none	0.7665	0.1187	0.9244	0.7690	0.6993	0.7002
You et al.[18]	YouDMOS_p	VSNR	0.8738	0.4102	0.9087	0.8866	*0.7859*	0.8738
	QUALITAS	none	**0.9467**	**0.7557**	**0.9333**	**0.9397**	**0.8684**	**0.9392**

Table VII. Performance across SIQA-SET in predicting perceived stereoscopic image quality: Spearman's Rank Ordered Correlation Coefficient (SROCC). Bold indicates the best metric, while italics indicate the second best.

Author	SIQA	NIQA	JP2K	JPEG	WN	Blur	FF	ALL
Benoit et al.[5]	d_1	UQI	*0.9104*	*0.7268*	0.9248	*0.9306*	*0.8162*	**0.9335**
Gorley et al.[8]	SBLC	none	0.6744	0.4431	0.6219	0.6229	0.2133	0.5963
Gu et al.[9]	ODDM_4	none	0.8131	0.4202	0.9206	0.6577	0.7734	0.7223
Hewage et al.[10]	$\text{PSNR}_{\text{edge}}$	VIFP	0.7802	0.2360	0.8616	0.7958	0.5027	0.7976
Jin et al.[11]	MSE_{ms}	AD	0.8608	0.4484	0.9310	**0.9318**	0.6859	0.8952
Joveluro et al.[12]	PQM_{3D}	none	0.0239	0.1329	0.9167	0.1398	0.3360	0.2667
Mao et al.[13]	Q_{mao}	none	0.7460	0.1629	0.7790	0.6279	0.3599	0.7253
Yang et al.[17]	IQA	none	0.7993	0.1212	0.9316	0.9020	0.5875	0.8340
You et al.[18]	YouDMOS_p	MSSIM	0.8979	0.5991	**0.9423**	0.9282	0.7349	0.9223
	QUALITAS	none	**0.9126**	**0.7384**	*0.9336*	0.9290	**0.8256**	*0.9334*

into account certain distortions. Therefore, we propose an analysis of the behavior of the top ten metrics in the overall SIQA-SET, in Tables VI–IX. Each table represents only one performance measure, LCC, SROCC, KROCC or RMSE.

From Table VI, QUALITAS obtained the best results in linear correlation coefficient not only in overall performance (93.92%) but also in all individual distortions. $d_1 \leftarrow$ FSIM obtained the second best results in linear correlation coefficient not only in overall performance (91.69%) but also in all individual distortions, except in fast fading (FF) distortion. For FF distortion, YouDMOS_p is the second best metric, with 78.59%.

Table VII shows the performance across SIQA-SET including QUALITAS in predicting perceived stereoscopic image quality using Spearman's Rank Ordered Correlation Coefficient. Here, $d_2 \leftarrow$ UQI obtained the best results in overall performance (93.35%) and the second best results in JPEG2000 (91.04%), JPEG (62.59%), Gaussian blur (93.06%) and fast fading distortions (81.62%). Meanwhile, QUALITAS obtained the second best results in overall performance (93.34%) and the best results in JPEG2000 (91.26%), JPEG (73.84%) and fast fading distortions (82.56%). For Gaussian blur (93.06%) and white noise (93.18%), the best metrics are YouDMOS_p and MSE_{ms}, respectively.

From Table VIII, $A_v \leftarrow$ MAD obtained the best results in linear correlation coefficient not only in overall performance (77.72%) but also in all individual distortions, except in fast fading distortion. For fast fading distortion, $\text{YouDMOS}_p \leftarrow$ UQI is the best metric, with 64.47%.

Table IX shows the performance across SIQA-SET in predicting perceived stereoscopic image quality using Root Mean Squared Error. Here, $A_v \leftarrow$ MAD obtained the best results not only in overall performance (0.0732) but also in JPEG2000 (0.0630), JPEG (0.0529), white noise (0.0805), Gaussian blur (0.0919) and fast fading distortion (0.0861).

CONCLUSIONS

In this article we propose a new stereoscopic image quality metric, QUALITAS, which works in the following way.

- By dividing the disparity map into two parts.

Table VIII. Performance across SIQA-SET in predicting perceived stereoscopic image quality: Kendall's Rank Ordered Correlation Coefficient (KROCC). Bold indicates the best metric, while italics indicate the second best.

Author	SIQA	NIQA	JP2K	JPEG	WN	Blur	FF	ALL
Benoit et al.[5]	d_2	UQI	*0.7405*	*0.5212*	0.7570	0.7697	*0.6245*	*0.7659*
Gorley et al.[8]	SBLC	none	0.4608	0.3065	0.4468	0.4141	0.1510	0.4201
Gu et al.[9]	$ODDM_4$	none	0.6089	0.2666	0.7456	0.5071	0.5783	0.5284
Hewage et al.[10]	$PSNR_{edge}$	VIFP	0.5899	0.1596	0.6652	0.6222	0.3694	0.5958
Jin et al.[11]	MSE_{ms}	AD	0.6620	0.2869	0.7665	*0.7717*	0.4998	0.7022
Joveluro et al.[12]	PQM_{3D}	none	0.0152	0.0893	0.7473	0.0929	0.2263	0.1869
Mao et al.[13]	Q_{mao}	none	0.5418	0.1216	0.5791	0.4525	0.2535	0.5294
Yang et al.[17]	IQA	none	0.5918	0.0735	0.7665	0.7333	0.4168	0.6296
You et al.[18]	$YouDMOS_p$	MSSIM	0.7158	0.4066	**0.7905**	**0.7737**	0.5485	0.7462
	QUALITAS	none	**0.7443**	**0.5396**	*0.7703*	0.7697	**0.6371**	**0.7668**

Table IX. Performance across SIQA-SET in predicting perceived stereoscopic image quality: Root Mean Squared Error (RMSE). Bold indicates the best metric, while italics indicate the second best.

Author	SIQA	NIQA	JP2K	JPEG	WN	Blur	FF	ALL
Benoit et al.[5]	d_1	BRISQUE	0.1935	0.1750	**0.1001**	*0.1442*	**0.1278**	0.1485
Gorley et al.[8]	SBLC	none	0.2980	0.1923	0.4733	0.3872	0.5295	0.3750
Gu et al.[9]	$ODDM_4$	none	*0.1041*	0.1654	*0.1079*	**0.1304**	0.1490	*0.1315*
Hewage et al.[10]	$PSNR_{edge}$	AD	0.1995	0.1875	0.2949	0.2203	*0.1449*	0.2084
Jin et al.[11]	MSE_{dp}	BIQI	0.1840	0.1726	0.1130	0.2714	0.1778	0.1754
Mao et al.[13]	Q_{mao}	none	0.3899	0.4327	0.3921	0.3730	0.3014	0.3783
Shen et al.[15]	HDPSNR	none	0.2035	0.2237	0.2981	0.2162	0.2549	0.2415
Yang et al.[17]	IQA	none	0.2081	*0.0931*	0.4136	0.3473	0.4275	0.2932
You et al.[18]	DQ_{map2}	none	**0.0961**	**0.0742**	0.1273	0.2506	0.1498	**0.1289**
	QUALITAS	none	0.4907	0.5838	0.4539	0.4799	0.3705	0.4754

- By employing a Contrast Band-Pass Filtering, so that dynamic parameters are considered as observational distances.
- By including three factors: loss of correlation, luminance and contrast distortion.
- By taking into account the visual differences between left and right images, employing a penalization depending on their wavelet energy.

Thus, the novelty of QUALITAS lies in combining some features of certain stereoscopic image quality assessments.

Furthermore, this article includes the comparison of 25 Stereoscopic Image Quality Assessments (SIQAs).

Some algorithms can be combined with any 2D/normal quality metric (NIQA), giving as a result 221 metrics which were compared against QUALITAS. QUALITAS obtained the best results in terms of overall performance of the correlation coefficients LCC, SROCC and KROCC, with 93.92%, 93.34% (just 0.01% below the best one) and 76.68%. For Root Mean Squared Error, QUALITAS scored in the lower rank, but it is well known that RMSE is not correlated with human vision.

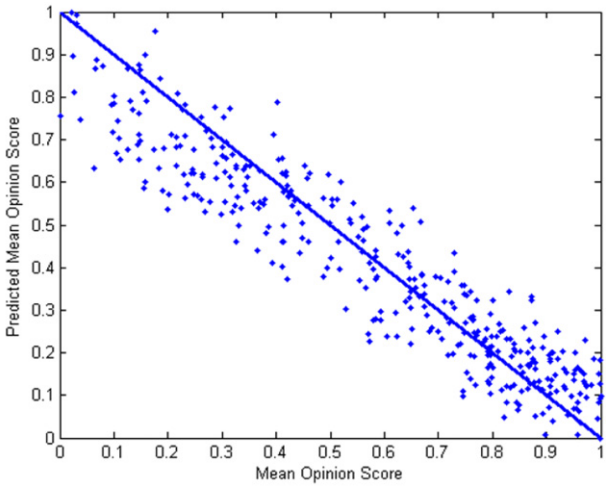


Figure 11. *MOS* versus MOS_p (both normalized). MOS_p is predicted by QUALITAS.

These results confirm that all metrics in SIQA-SET are simple modifications of NIQA, which take in to account

some extra characteristics from the disparity map (usually depth variances). On the contrary, QUALITAS incorporates disparity masking in addition to divide the 3D scenario into two parts: background and foreground planes.

Some distortions considered, such as additive white Gaussian noise and Gaussian blur, are global distortions. Therefore, these distortions do not affect the perception of depth considerably but they statistically modify the image content. Some metrics in SIQA-SET correlate extremely well with these distortions, such as d_1 and MSE_{ms}.

For distortions with localized artifacts, the performance of all metrics in SIQA-SET decreases, especially for the local blocking artifacts caused by a JPEG compression. Thus, for JPEG compression distortion, the performance of QUALITAS is good, with 75.57% (LCC), showing to be not dependent on a monoscopic image quality.

Finally, the presented tests and their results globally confirm the potential of the proposed metric to evaluate stereo-pairs.

ACKNOWLEDGMENT

This work is supported by the National Polytechnic Institute of Mexico by means of Project No. 20140096 granted by the Academic Secretary, National Council of Science and Technology of Mexico by means of Project No. 204151/2013, LABEX Σ-LIM France, Coimbra Group Scholarship Programme granted by University of Poitiers and Region Poitou-Charentes (France), and Project Smart-k, MIUR—Regione Lombardia (Italy).

REFERENCES

[1] A. K. Moorthy, C. C. Su, A. Mittal, and A. C. Bovik, "Subjective evaluation of stereoscopic image quality," Signal Process., Image Commun. **28**, 3379 (2013).

[2] L. Goldmann, F. De Simone, and T. Ebrahimi, "Impact of acquisition distortion on the quality of stereoscopic images," *Proc. Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2010).

[3] X. Wang, M. Yu, Y. Yang, and G. Jiang, "Research on subjective stereoscopic image quality assessment," Proc. SPIE **7255**, 725509 (2009).

[4] R. Akhter, Z. M. Parvez Sazzad, Y. Horita, and J. Baltes, "No-reference stereoscopic image quality assessment," Proc. SPIE **7224**, 75240 (2010).

[5] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," EURASIP J. Image Video Process. 659024 (2008).

[6] E. Bosc, R. Pepion, P. Le Callet, M. Koppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, "Towards a new quality metric for 3-d synthesized view assessment," IEEE J. Sel. Top. Signal Process. **5**, 1332 (2011).

[7] M. J. Chen, C. C. Su, D. K. Kwon, L. K. Cormack, and A. C. Bovik, "Full-reference quality assessment of stereoscopic images by modeling binocular rivalry," *46th Annual Asilomar Conf. on Signals, Systems, and Computers* (IEEE, New York, 2012).

[8] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," Proc. SPIE **6803**, 680305 (2008).

[9] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new no-reference stereoscopic image quality assessment based on ocular dominance theory and degree of parallax," *21st Int'l. Conf. on Pattern Recognition (ICPR)* (IEEE, New York, 2012), pp. 206–209.

[10] C. Hewage and M. Martini, "Reduced-reference quality metric for 3d depth map transmission," *3DTV-Conf. The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON)* (IEEE, New York, 2010).

[11] L. Jin, A. Boev, A. Gotchev, and K. Egiazarian, "3d-dct based multiscale full-reference quality metric for stereoscopic video," *Proc. 6th Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2012).

[12] P. Joveluro, H. Malekmohamadi, W. A. C. Fernando, and A. Kondoz, "Perceptual video quality metric for 3d video quality assessment," *3DTV-Conf. The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON)* (IEEE, New York, 2010).

[13] X. Mao, M. Yu, X. Wang, G. Jiang, Z. Peng, and J. Zhou, "Stereoscopic image quality assessment model with three-component weighted structure similarity," *Int'l. Conf. on Audio Language and Image Processing (ICALIP)* (IEEE, New York, 2010), pp. 1175–1179.

[14] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," IEEE Trans. Image Process. **22**, 1940 (2013).

[15] L. Shen, J. Yang, and Z. Zhang, "Stereo picture quality estimation based on a multiple channel hvs model," *2nd Int'l. Congress on Image and Signal Processing (CISP)* (IEEE, New York, 2009).

[16] M. Solh, J. M. Bauza, and G. AlRegib, "3VQM: a vision-based quality measure for DIBR-based 3D videos," *2011 IEEE Int'l. Conf. on Multimedia and Expo.* (IEEE, Piscataway, NJ, 2011), pp. 1–6.

[17] J. Yang, C. Hou, Y. Zhou, Z. Zhang, and J. Guo, "Objective quality assessment method of stereo images," *3DTV-Conf. The True Vision—Capture, Transmission and Display of 3D Video* (IEEE, New York, 2009), pp. 9–12.

[18] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis," *Proc. Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2010).

[19] Z. Zhu and Y. Wang, "Perceptual distortion metric for stereo video quality evaluation," WSEAS Trans. Signal Process. **5**, 241 (2009).

[20] C. U. Visual Communications Laboratory website, MeTriX MuX visual quality assessment package, http://foulard.ece.cornell.edu/gaubatz/metrix_mux/. Cornell University Visual Communications Laboratory, accessed 2014.

[21] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," *Conf. Record of the Thirty-Seventh Asilomar Conf. on Signals, Systems and Computers* (IEEE, New York, 2003), Vol. 2, pp. 1398–1402.

[22] D. Chandler and S. Hemami, "Vsnr: a wavelet-based visual signal-to-noise ratio for natural images," IEEE Trans. Image Process. **16**, 2284 (2007).

[23] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600 (2004).

[24] H. Sheikh and A. Bovik, "Image information and visual quality," IEEE Trans. Image Process. **15**, 430 (2006).

[25] Z. Wang and A. Bovik, "A universal image quality index," IEEE Signal Process. Lett. **9**, 81 (2002).

[26] H. Sheikh, A. Bovik, and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," IEEE Trans. Image Process. **14**, 1918 (2005).

[27] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A. Bovik, "Image quality assessment based on a degradation model," IEEE Trans. Image Process. **9**, 636 (2000).

[28] T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for formulation of quality measures incorporated in halftoning algorithms," *IEEE Int'l. Conf. on Acustics, Speech and Signal Processing* (IEEE, New York, 1993), Vol. 5, p. 301.

[29] A. Moorthy and A. Bovik, "A two-step framework for constructing blind image quality indices," IEEE Signal Process. Lett. **17**, 513 (2010).

[30] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Trans. Image Process. **21**, 4695 (2012).

[31] A. Mittal, R. Soundararajan, and A. Bovik, "Making a 'completely blind' image quality analyzer," IEEE Signal Process. Lett. **20**, 209 (2013).

[32] J. Moreno, B. Jaime, and C. Fernandez-Maloigne, "NRPSNR: no-reference peak signal-to-noise ratio for JPEG2000," *Data Compression Conf. (DCC)* (Snowbird, Utah, 2013), pp. 511–511.

[33] J. Moreno, "P2SNR: perceptual full-reference image quality assessment for JPEG2000," *Data Compression Conf. (DCC)* (Snowbird, Utah, 2012), pp. 406–406.
[34] L. Zhang, D. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," IEEE Trans. Image Process. **20**, 2378 (2011).
[35] L. Zhang, D. Zhang, and X. Mou, "Rfsim: a feature based image quality assessment metric using Riesz transforms," *17th IEEE Int'l. Conf. on Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2010), pp. 321–324.
[36] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "Two new full-reference quality metrics based on HVS," *Proc. 2nd Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2006), p. 4.
[37] Z. Wang, H. R. Sheikh, and A. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," *Int'l. Conf. on Image Processing* (IEEE, New York, 2002), Vol. 1, pp. I–477–I–480.
[38] F. Zhang, S. Li, L. Ma, and K. N. Ngan, "Limitation and challenges of image quality measurement," *Visual Communications and Image Processing (International Society for Optics and Photonics)* (SPIE, Bellingham, Washington, 2010), p. 774402.
[39] 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).
[40] ITU, Bt-500-11: Methodology for the Subjective Assessment of the Quality of Television Pictures, 2002.
[41] M. Park, J. Luo, and A. Gallagher, "Toward assessing and improving the quality of stereo images," IEEE J. Sel. Top. Signal Process. **6**, 460 (2012).
[42] M. Huan, A. Minazuki, and H. Hayashi, "Study on 3d image assessment using motion capture system," *Int'l. Conf. Advanced Applied Informatics (IIAIAAI)* (IEEE, New York, 2012), pp. 182–186.
[43] X. Otazu, C. Párraga, and M. Vanrell, "Toward a unified chromatic induction model," J. Vis. **10** (2010).
[44] L. B. Stelmach and W. J. Tam, "Stereoscopic image coding: effect of disparate image-quality in left- and right-eye views," Signal Process, Image Commun. **14**, 111 (1998).
[45] R. Palaniappan and N. Jayant, "Subjective quality in 3dtv: effects of unequal bit allocation to left and right views," *6th Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2012).
[46] D. Vatolin, M. Erofeev, A. Zachesov, D. Sumin, D. Akimov, and A. Fedorov, "Testing methods for 3d content viewing devices," *6th Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics* (Arizona State University, Arizona, 2012).
[47] M.-J. Chen, D.-K. Kwon, A. C. Bovik, and L. K. Cormack, "Optimizing 3d image display using the stereoacuity function," *IEEE Int'l. Conf. on Image Processing* (IEEE, Piscataway, NJ, 2012).
[48] G. van de Wouwer, P. Scheunders, and D. van Dyck, "Statistical texture characterization from discrete wavelet representations," IEEE Trans. Image Process. **8**, 592 (1999).
[49] B. A. Wilson and M. A. Bayoumi, "A computational kernel for fast and efficient compressed-domain calculations of wavelet subband energies," IEEE Trans. Circuits Syst. II **50**, 389 (2003).
[50] S. Yasakethu, D. De Silva, W. Fernando, and A. Kondoz, "Predicting sensation of depth in 3d video," Electron. Lett. **46**, 837 (2010).
[51] S. L. P. Yasakethu, S. Worrall, D. V. S. X. De Silva, W. A. C. Fernando, and A. Kondoz, "A compound depth and image quality metric for measuring the effects of packet loss on 3d video," *17th Int'l. Conf. on Digital Signal Processing* (IEEE, New York, 2011), pp. 1–7.