




Article

Objective Prediction of Human Visual Acuity Using Image Quality Metrics

Julián Espinosa Tomás , Jorge Pérez Rodríguez, David Más Candela , Carmen Vázquez Ferri and Esther Perales * 

Departamento de Óptica Farmacología y Anatomía, Universidad de Alicante, 03690 Alicante, Spain; julian.espinosa@ua.es (J.E.T.); jorge.perez@ua.es (J.P.R.); david.mas@ua.es (D.M.C.); ferri@ua.es (C.V.F.)

* Correspondence: esther.perales@ua.es

Featured Application: It is often difficult to subjectively determine subjects' uncorrected visual acuity given their age, lack of cooperation with subjective measurement, etc. In other cases, it would be desirable to be able to predict this visual acuity prior to any type of ocular intervention. The proposed method allows such a determination to be objectively made by determining the degradation of the eye's optical system from a set of natural images.

Abstract: This work addresses the objective prediction of human uncorrected decimal visual acuity, an unsolved challenge due to the contribution of both physical and neural factors. An alternative approach to assess the image quality of the human visual system can be addressed from the image and video processing perspective. Human tolerance to image degradation is quantified by mean opinion scores, and several image quality assessment algorithms are used to maintain, control, and improve the quality of processed images. The aberration map of the eye is used to obtain the degraded theoretical image from a set of natural images. The amount of distortion added by the eye to the natural image was quantified using different image processing metrics, and the correlation between the result of each metric and subjective visual acuity was assessed. The correlation obtained for a model based on a linear combination of the normalized mean square error metric and the feature similarity index metric was very good. It was concluded that the proposed method could be an objective way to determine subjects' monocular and uncorrected decimal visual acuity with low uncertainty.

Keywords: visual acuity; aberration; image quality assessment



Citation: Tomás, J.E.; Rodríguez, J.P.; Candela, D.M.; Ferri, C.V.; Perales, E. Objective Prediction of Human Visual Acuity Using Image Quality Metrics. *Appl. Sci.* **2023**, *13*, 6350. <https://doi.org/10.3390/app13106350>

Academic Editor: Hariton-Nicolae Costin

Received: 3 April 2023
Revised: 16 May 2023
Accepted: 19 May 2023
Published: 22 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The quality of the optical system, the quality of the retinal image, and subjective visual quality are three highly related concepts in the visual optics field. They are approached by objective functions, such as the modulation transfer function (MTF) of the system, the Zernike decomposition of the wavefront, or by subjective parameters, such as subjects' decimal uncorrected visual acuity (VA). The relation among these three concepts is clear: if the quality of the human visual system (HVS) is poor, the quality of the resulting image will also be bad and, as a result, VA will be poor.

The retinal image is affected by aberrations of the system, scattering and diffraction of light, and retinal sampling. Nevertheless, the vision process is not simple because it results from the proper combination of physical, optical, physiological, neural, and psychological aspects. Thus, it is relatively normal to find people who indicate they "correctly view" a certain image because they recognize the structure of the object, but not its details. Therefore, they pose no need for refractive correction. Guirao and Williams [1] suggested that the visual quality metrics obtained on the retinal plane are more consistent with subjective measurements than those calculated on the pupil plane. Presently, the most

widespread objective criterion to predict visual quality is the visual Strehl ratio (*VSOTF*). Cheng et al. [2] obtained a correlation between defocus and astigmatism, and 31 different visual quality metrics. They concluded that the *VSOTF* might be a good objective parameter. They also concluded that the value of the root mean square (*RMS*) wavefront error, or other parameters like the Strehl ratio (*SR*), are not reliable indicators of the subjective quality of the retinal image. Thus, they chose a metric that combined the point spread function (*PSF*) of the system with a spatial sensitivity function. Marsack et al. [3] demonstrated the need for single-value metrics other than *RMS* to assess the VA effects of low aberration levels. Later, Watson and Ahumada [4] proposed a model for VA that incorporates the set of ocular aberrations, optical and neural filtering, and neural noise.

An alternative approach to assess the image quality of the HVS can be addressed from the perspective of the image and video processing field. On the one hand, human tolerance to image degradation is quantified with mean opinion scores (*MOS*) [5]. On the other hand, a series of image quality assessment (*IQA*) algorithms are used to maintain, control, and improve the quality of processed images. Early works focused on comparing the degraded image to the initial one by using two objective parameters: peak signal-to-noise ratio (*PSNR*) and mean square error (*MSE*) or its square root (*RMSE*). *MSE* is the standard method applied in image comparisons because it is simple and fast but is not usually a good estimator of subjective perception because it does not consider HVS characteristics. It is also an unbounded metric, which makes it difficult to correlate with VA [6]. The *PSNR* metric is also based on pixel-by-pixel by comparing the reference image to the distorted image through *MSE*, and it is still one of the most popular ways to assess the quality difference between images [7]. Its disadvantages are that it is not a bounded metric, and it does not consider HVS properties. Thus, it does not correlate well with subjective tests. In the last few years, one of the most commonly used metrics has been the structural similarity index (*SSIM*) [8] because of the good correlation with *MOS*.

IQA objective methods can be classified into three types, full-reference, reduced-reference, and non-reference, depending on the use of a reference image, some of the information of that image, or no available reference image to make the comparison, respectively. In this paper, we focus on studying seven metrics that derive from full-reference algorithms to objectively determine subjects' VA. They are *MSE*, *PSNR*, *SSIM*, the Multi-scale Structural Similarity Index (*MSSSIM*) [9], the peak signal-to-noise ratio based on the HVS (*PSNR-HVS*) [10], gradient magnitude similarity deviation (*GMSD*) [11] and the Feature Similarity Index (*FSIM*) [12]. These metrics were chosen because they are fast, easy to implement, and sufficiently confirmed.

We establish a relation between image quality metrics, hitherto restricted to the field of signal and image processing, and the quality of vision concept, quantified by monocular VA. As far as we know, no works use static image or video quality metrics to study the quality of human vision, and only two references that link both aspects appear in the literature. Iskander [13] discussed a possible relation between some image processing metrics and subjects' visual quality. The study was performed for two sets of metrics: those based on comparing images and those based on the optical transfer function. The metric that best correlated with the evaluation of subjects' ametropia was entropy. Later in [14], some of the authors adapted the *MSSIM* metric to the visual process (*VMSSIM*) with two subjects (one myopic and another hyperopic), both before and after being treated by LASIK surgery, to objectively predict their visual quality.

Several studies have been recently published on quantifying the quality of an image, but by combining several full reference metrics [15–17]. Their results reveal that these combinations seem to exceed individual metrics when predicting the quality of an image. We approached the problem of assessing HVS quality from the same point of view and proposed combining metrics to predict VA. Most psychophysical experiments are performed with relatively simple patterns, such as blobs, sinusoidal bars or grids, letters, etc. For example, the contrast sensitivity function is usually obtained from thresholds with global sinusoidal images. However, all these patterns are simpler than real-world images,

which can be considered as a superposition of a larger number of simple patterns. VA measures in some ways the degradation of a subject’s optimal visual quality. The objective of this work is to relate digital image processing metrics, that use natural images, with VA, since there must be a relationship between them. Subjective VA tests measured using optotypes comprises both physical and neural factors. Natural images are necessary to quantify in some manner for human tolerance to image degradation because using optotypes as images (not natural images) and the wavefront aberration (only physical factors) would not be able to model neural or even subjective factors that in some manners are present in natural images evaluated with *MOS*.

The manuscript is structured as follows. Section 2 describes the procedure. First, an image database is defined from existing ones. Next, details of the subjects participating in the study and measurements are provided. Then, the calculation of the PSF and the image of the eye are explained. The method section ends by describing the metrics calculation and defining the fitting to subjective VA and conditions. The results appear in the third section and the conclusions are finally stated. The metrics are described in Appendix A.

2. Materials and Methods

This section describes the method proposed to obtain an objective evaluation of subjects’ VA based on the application of image processing metrics and the physical data of eyes. Figure 1 shows the flow chart of the whole process. For a distant object, a hyperopic eye is supposed to use crystalline lens accommodation to focus the image on the retina and thus, obtaining the maximum quality of vision. Such accommodation changes the value of the Zernike coefficient related to the Z_2^0 [18] following:

$$Z_2^0 = \begin{cases} R_p^2 \frac{(S-Ac) + \frac{C}{2}}{-4\sqrt{3}}; & \text{if } S \geq Ac \\ R_p^2 \frac{\frac{C}{2}}{-4\sqrt{3}}; & \text{if } S < Ac \end{cases} \quad (1)$$

where *S* and *C* are the values, in diopters, of the sphere and cylinder of the studied eye. Thus, to obtain *PSF* in hyperopic subjects, the possible accommodation of the eye was considered through the adjustment of the average monocular accommodation of a subject as a function of age provided by Duane [19] to a third-degree polynomial.

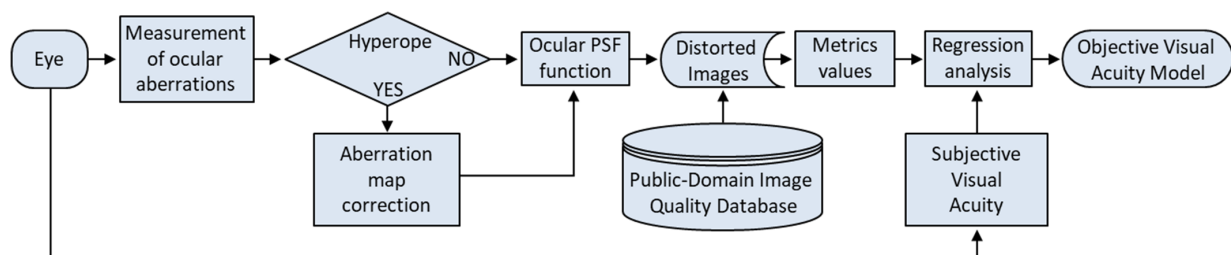


Figure 1. Flow chart of the proposed approach.

2.1. Image Database

The image databases commonly used in image processing are available on the World Wide Web. They are composed of reference natural images, the corresponding degraded images, and the mean *MOS* and/or differential mean opinion score (*DMOS*) values for a significant number of subjects. We randomly chose a set of 49 different reference images that belong to three distinct image databases: 29 to the LIVE base from the image and video engineering laboratory of the University of Texas [20]; 10 from the *IVC* database of the research group of communication of images and video of the Research Institute on Communications and Cybernetics [21]; the last 10 images belong to the Toyama-MICT base [22].

2.2. Subjects

We studied 52 randomly selected eyes of 52 subjects of both sexes (50% women, 50% men) who had not suffered any eye disease or trauma. Their age range was wide (18 to 62 years old). The study did not present any invasive action. The tests to perform, their nature, and their purpose were explained to all the subjects. They agreed to undergo them and provided their consent. Experiments were conducted with the approval of the Ethics Committee of the University of Alicante and in accordance with the Declaration of Helsinki. Non-cycloplegic subjective test refractions and subjective monocular logMAR VA without correction for distant vision under photopic lighting conditions (85 cd/m^2) were conducted by optometrists. Visionix VX-120 was used to capture corneal topographies by Placido's ring-based technology, aberration maps of the eye, and tonometry, pachymetry, and anterior chamber data. All these measurements were taken three times per subject during sessions separated by a 24 h time interval. LogMAR VA values were converted to decimal VA for convenience reasons following the relation $VA_{decimal} = 10^{(-\log MAR)}$. Pupillary diameters were not measured during the VA assessment. Photopic conditions establish natural pupil diameters ranging between 2 and 4 mm [23] and it is assumed that VA does not vary within that pupil diameter range.

In Figure 2, we represent the characteristics of the studied eyes. Figure 2A shows the refractive errors associated with each examined age group. The prevalence of nearsightedness over farsightedness, with a maximum spherical equivalent refraction of 0.75 D, is evident in most subjects. Figure 2B illustrates the average spherical equivalent of the eyes associated with each age range. Finally, Figure 2C depicts a histogram of the number of eyes with different subjective VA ranges.

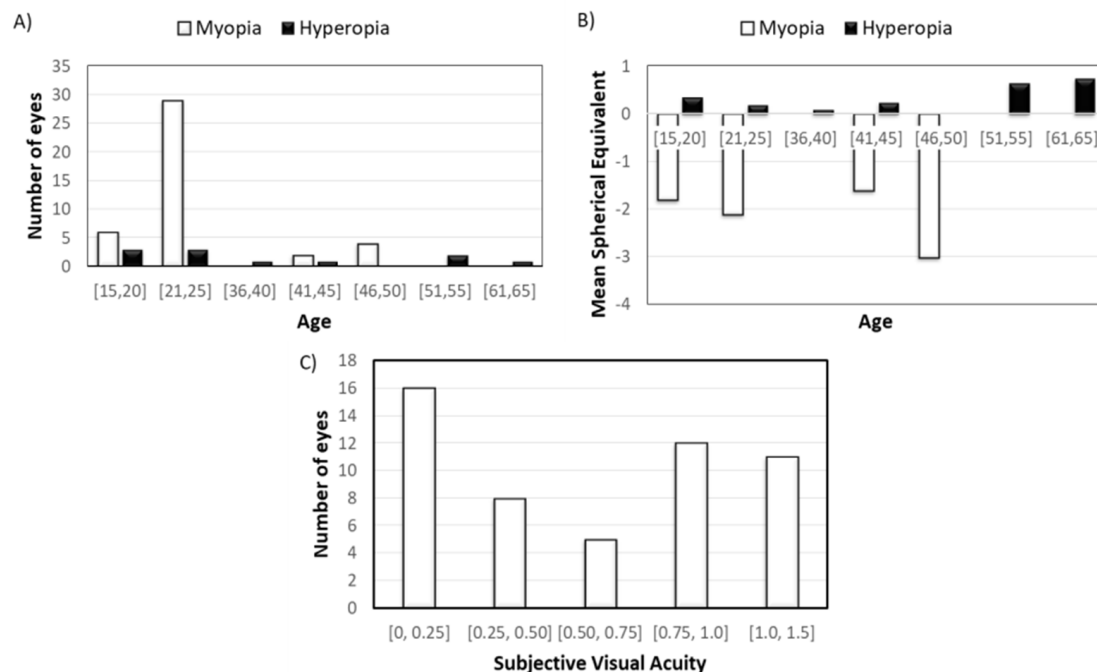


Figure 2. Characteristics of the studied eyes. (A) Type of refractive error (myopia-hyperopia). (B) Spherical equivalent associated with age range. (C) Number of subjects associated with the subjective decimal VA ranges.

2.3. The Point Spread Function and the Image of the Eye

The monochromatic PSF of the subject was computed from the wave aberration function $W(x, y)$ reconstructed from the Zernike coefficients measured with Visionix VX-120. In this calculation, the first three Zernike coefficients corresponding to the aberration components of piston, tilt X and tilt Y were not considered because they constitute translations

and tilts of the reference system, which can be naturally compensated with eye movements. The *PSF* is obtained as the squared module of the Fourier transform of the generalized pupil function [24]:

$$PSF(x, y) = |FT(P(x, y))|^2 \tag{2}$$

We considered the Stiles–Crawford effect due to the anatomical structure of photoreceptors [25], which can be modeled by an apodizing filter located at the entrance pupil [26]. Then, the generalized pupil function is given by [27]:

$$P(x, y) = e^{-0.116 R_p^2(x^2+y^2)} e^{-ikW(x,y)}, \tag{3}$$

where R_p is the pupillary radius, which was provided by the Visionix VX-120 system. In the work of Prakash et al. [28], it is shown that under photopic conditions the pooled pupil diameter is 4.07 ± 0.63 mm so the Zernike expansion coefficients provided by Visionix were transformed into new ones [29,30] for pupil diameters above 4.07 mm.

Finally, we considered that the eye’s image of an object could be obtained through the convolution of the function that represents the object $O(x, y)$ with the *PSF* of the optical system. By applying the convolution theorem [24], this image can be obtained as the inverse Fourier transform of the product of the convoluted functions, i.e.,

$$I(x, y) = \mathcal{F}^{-1}\{\mathcal{F}[O(x, y)] \cdot \mathcal{F}[PSF(x, y)]\} \tag{4}$$

2.4. The Point Spread Function and the Image of the Eye

Having defined how to obtain the degraded eye’s image of an object, the values of the seven studied metrics were determined. We established the metrics result for each eye as the average of the metrics values for the 49 object images.

Subjects’ subjective VA and the results obtained from calculating the metrics and linear combinations of metrics were fitted with a monotonous logistic function that is commonly used to study image quality [15]. It comes as follows:

$$VA(Q) = a \left(\frac{1}{2} - \frac{1}{1 + \exp[b(Q - c)]} \right) + dQ + e, \tag{5}$$

where Q represents any of the used metrics or a linear combination Q_L , and (a, b, c, d, e) are the parameters to determine.

The first issue to address in the proposal of the linear combinations of full reference metrics is collinearity. Collinearity is a regression analysis problem; if predictors are in a linear combination, the influence of each one on the criterion cannot be distinguished by overlapping them with one another. In that case, the confidence intervals of the estimating coefficients are often wide, which indicates that the obtained estimates are imprecise and probably unstable. The difficulty of assessing the existence of collinearity lies in determining the maximum degree of the permissible relation between independent variables. No consensus on this issue has been reached. Neter et al. [31] considered a series of indicators to analyze the degree of multicollinearity among the regressors of a multivariate linear model. The simplest is the variance inflation factor (*VIF*) between two of the regressive variables, which is defined as:

$$VIF = \frac{1}{1 - R^2}, \tag{6}$$

where R^2 is the coefficient of determination between the two variables. According to these authors, if *VIF* is higher than 10, it can be concluded that the collinearity between the two selected variables is high and will affect multilinear fit by increasing the variance value.

Implementing regression fits requires compliance with a series of assumptions to reach conclusive results. These are the homoscedasticity, normality, and independence of residuals. The homoscedasticity hypothesis establishes that the variability of residuals is independent of the explanatory variables. Failure to comply with this condition may

result in the fitted parameters varying according to sample size. Regarding normality, residuals should follow a normal distribution with a zero average. Last but not least, the Durbin–Watson (*DW*) [32] statistic can be used to verify the independence of residuals, r_i :

$$DW = \frac{\sum_{i=2}^n (r_i - r_{i-1})^2}{\sum_{i=1}^n r_i^2}; \tag{7}$$

where n is the number of eyes. The null hypothesis (no statistical evidence that residuals are positively self-correlate) is rejected if the *DW* value is less than a lower critical value $DW_{(L,\alpha)}$, where α represents the significance level (in this work, $\alpha = 0.05$). If *DW* is higher than an upper $DW_{(U,\alpha)}$ critical value, it is accepted that there is no correlation. In intermediate cases, the test is not conclusive. If the value $(4 - DW)$ is less than $DW_{(L,\alpha)}$, there is statistical evidence that residuals negatively self-correlate. If the value $(4 - DW)$ is higher than $DW_{(U,\alpha)}$, there is no statistical evidence for a negative self-autocorrelation [33].

Besides fulfilling the above assumptions, it is necessary to establish a criterion to select the metric model that best predicts VA. Wei et al. [34] introduced an objective index (*WI*) to evaluate the performance of an estimation model. In general, the more accurate the model is, the bigger *R* is and the smaller the mean of the squares of residuals. The *WI* index is defined as:

$$WI = \frac{R}{\frac{1}{n} \sum_{i=1}^n r_i^2}; \tag{8}$$

Another descriptive statistic, which measures the dispersion of a dataset and can be used to compare models' performance, is the quartile coefficient of dispersion (*QCD*) [35]. It is defined by:

$$\overline{QCD} = \frac{1}{n} \sum_{i=1}^n \left(\frac{Q_3 - Q_1}{Q_3 + Q_1} \right)_i; \tag{9}$$

where Q_1 and Q_3 are the first and third quartiles of the dataset, which consists of the results of each metric of every eye over the set of 49 images. The higher this coefficient is, the wider the data variability is. Therefore, a model with a low \overline{QCD} is expected.

3. Results and Discussion

The 49 mathematically degraded images that resulted from the convolution of the *PSF* of each eye with the original images were obtained. Subsequently, the values of all the metrics for each degraded image, their mean, and their standard deviation (*SD*) were computed.

Due to the major difference between the values of the metrics, and to compare the coefficients that resulted from the distinct fittings, the metrics results were normalized within the range [0, 1]. We called them *nMSE*, *nPSNR*, *nGMSD*, and *nPSNR-HVS*. Therefore, linear combinations of *nMSE* with the other metrics were chosen, i.e., $Q_L = \beta_1 nMSE + \beta_2 Q$, with Q being any of the other six herein used metrics. First, we evaluated collinearity with the *VIF* index (Table 1). As shown, the *VIF* values were high, except for *nMSE*, which was related to the other metrics.

Table 1. VIF values among metrics.

	<i>nMSE</i>	<i>nPSNR</i>	<i>SSIM</i>	<i>nGMSD</i>	<i>MSSSIM</i>	<i>FSIM</i>
<i>nPSNR</i>	7.64					
<i>SSIM</i>	4.15	21.7				
<i>nGMSD</i>	3.05	12.3	27.8			
<i>MSSSIM</i>	4.23	24.8	150	30.8		
<i>FSIM</i>	5.93	35.8	76.7	15.2	57.4	
<i>nPSNR-HVS</i>	7.78	63.8	27.9	14.6	31.7	57.2

The generalized reduced gradient resolution algorithm was used to find the minimum value of the sum of the squares of the deviations between the value of subjective VA and

the target value provided by the fitting to the logistic function (5). This fitting provided an optimal local solution. Figures 3A–G and 4A–F show the subjective VA for distance vision versus the value obtained for all the metrics and the linear combination of metrics. The fitted functions are plotted in red, and the resulting parameters appear in Table 2. In all cases, the average amount of residuals was practically zero.

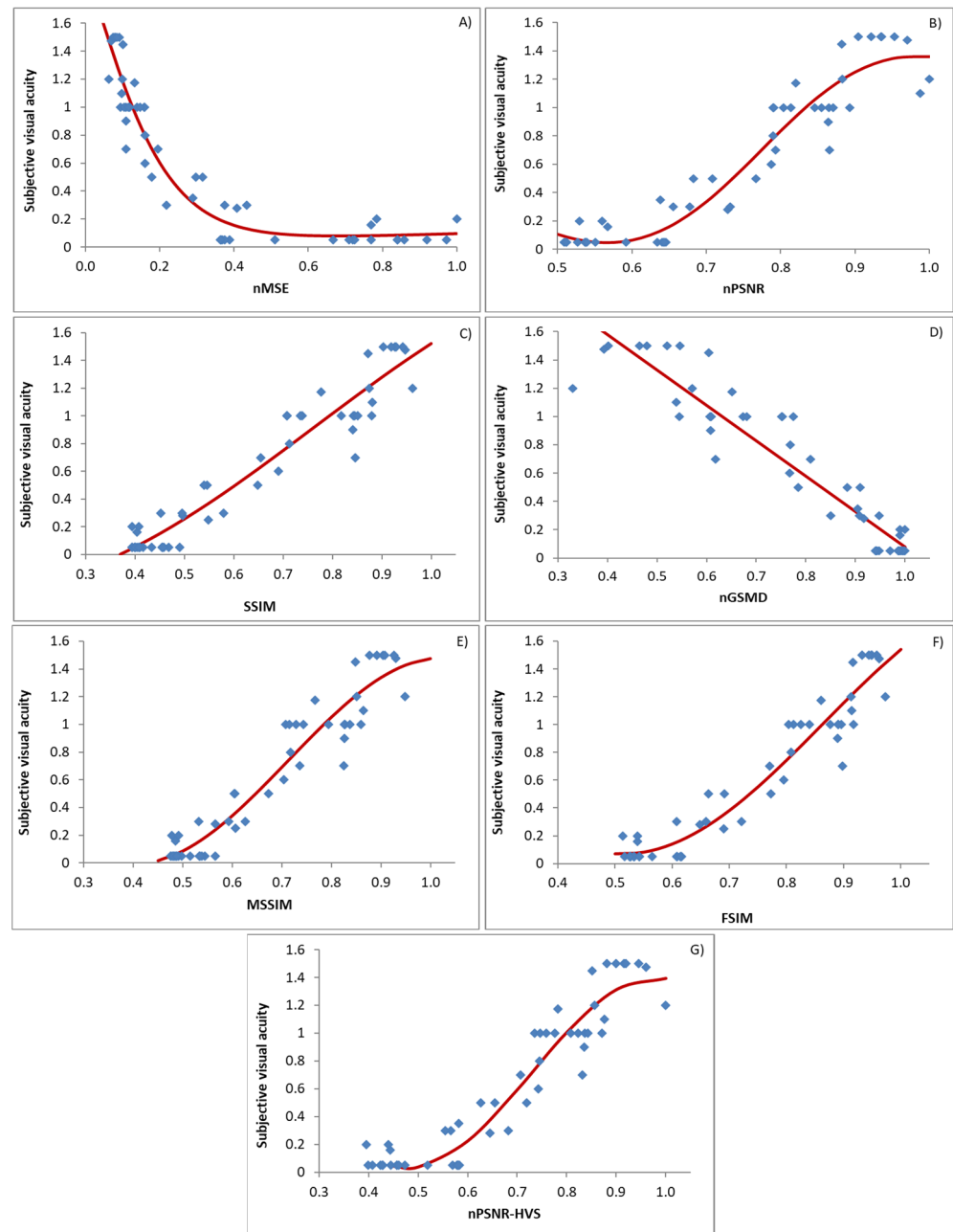


Figure 3. Subjective VA vs. the values from calculating the normalized single metrics. (A) *nMSE* (Mean Square Error Normalized), (B) *nPSNR* (Peak Signal-to-Noise Ratio Normalized), (C) *SSIM* (Structural Similarity Index), (D) *GMSD* (Gradient Magnitude Similarity Deviation), (E) *MSSIM* (Multiscale Structural Similarity Index), (F) *FSIM* (Feature Similarity Index), (G) *nPSNR-HVS* (Peak Signal-to-Noise Ratio based on the Human Visual System). Red lines represent the fitting to the logistic function (5).

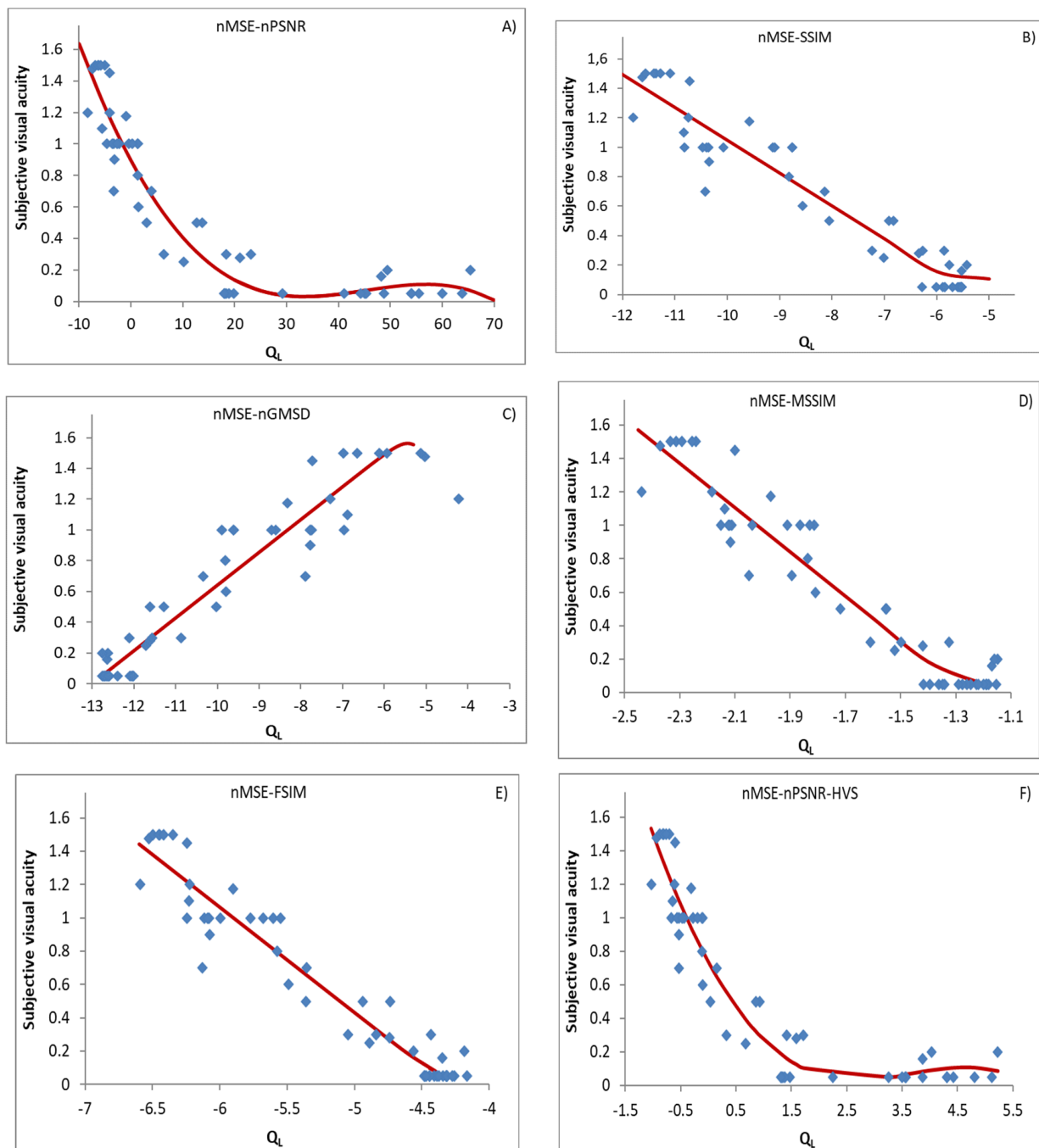


Figure 4. Subjective VA vs. the values from calculating the linear combinations of normalized single metrics. Q_L is the linear combination of $nMSE$ with the other metrics. (A) linear combination $nMSE$ (Mean Square Error Normalized) with $nPSNR$ (Peak Signal-to-Noise Ratio Normalized), (B) linear combination $nMSE$ with $SSIM$ (Structural Similarity Index), (C) linear combination $nMSE$ with $GMSD$ (Gradient Magnitude Similarity Deviation), (D) linear combination $nMSE$ with $MSSIM$ (Multiscale Structural Similarity Index), (E) linear combination $nMSE$ with $FSIM$ (Feature Similarity Index), (F) linear combination $nMSE$ with $nPSNR-HVS$ (Peak Signal-to-Noise Ratio based on the Human Visual System). Red lines represent the fitting to the logistic function (5).

Table 2. Parameters obtained for the fittings of the normalized metrics and the linear combinations of *nMSE* with the other metrics to the logistic function (5).

	a	b	c	d	e	β_1	β_2
<i>nMSE</i>	−4.000	8.992	0.000	0.078	2.016		
<i>nPSNR</i>	−4.364	−8.983	0.775	−4.522	4.211		
<i>SSIM</i>	10.29	2.228	0.746	−3.054	3.255		
<i>nGMSD</i>	0.002	0.000	−0.352	−2.501	2.580		
<i>MSSSIM</i>	24.57	3.075	0.712	−15.19	11.55		
<i>FSIM</i>	21.24	3.134	0.864	−12.47	11.78		
<i>nPSNR-HVS</i>	−23.08	−3.773	0.729	−17.55	13.51		
<i>nMSE-nPSNR</i>	64.78	0.021	45.16	−0.335	15.18	72.10	−12.75
<i>nMSE-SSIM</i>	0.173	155.1	−5.512	−0.223	−1.093	−0.772	−12.21
<i>nMSE-nGMSD</i>	−0.708	4.516	−4.870	0.213	2.418	−0.026	−12.78
<i>nMSE-MSSSIM</i>	37.89	17.24	−0.873	−1.326	17.27	0.111	−2.583
<i>nMSE-FSIM</i>	7.478	8.858	−3.770	−0.632	1.011	−0.944	−6.714
<i>nMSE-nPSNR-HVS</i>	41.54	0.282	3.670	−2.878	10.63	5.839	−1.391

To establish the best fitting, we performed an analysis of variance (ANOVA) with the subjective VA and the values provided by the above fittings of metrics. The ANOVA results are found in Table 3. In all cases, regressions provided a very low probability ($p < 0.001$) of accepting the null hypothesis. The F-number was used to determine whether the high coefficient of determination values occurred by chance. The critical value for a 95% confidence level and a number of points (eyes; $n = 52$) was 4.03. As the F-number is much higher than the critical value, a significant relationship between the variables in the model must be accepted. Therefore, all the metrics are useful for predicting the subjective VA value. Regarding the t statistic, the metrics provided t values above 20. The critical value was 2.01 with an alpha value equaling 0.05 (probability of the null hypothesis). As the obtained values were higher than the critical value, it can be stated that fittings were statistically significant with a probability over 95%. All the metrics also provided high coefficients of determination, which means that the logistic Function (5) allowed the studied variables to be correlated as a high percentage.

Table 3. ANOVA parameters for the above fittings. Coefficient of determination (R^2), standard error of estimate (σ_{est}), F-number, and t statistic.

	R^2	σ_{est}	F-Number	t Statistic
<i>nMSE</i>	0.9141	0.1563	532	23.1
<i>nPSNR</i>	0.9114	0.1588	514	22.7
<i>SSIM</i>	0.9218	0.1495	590	24.3
<i>nGMSD</i>	0.8891	0.1777	401	20.0
<i>MSSSIM</i>	0.9170	0.1540	553	23.5
<i>FSIM</i>	0.9266	0.1448	631	25.1
<i>nPSNR-HVS</i>	0.9152	0.1554	539	23.2
<i>nMSE-nPSNR</i>	0.9105	0.1600	509	22.6
<i>nMSE-SSIM</i>	0.9261	0.1454	626	25.0
<i>nMSE-nGMSD</i>	0.9178	0.1533	558	23.6
<i>nMSE-MSSSIM</i>	0.9228	0.1486	597	24.4
<i>nMSE-FSIM</i>	0.9309	0.1406	673	25.9
<i>nMSE-nPSNR-HVS</i>	0.9116	0.1589	516	22.7

The normality of the distribution of residuals was studied by applying three types of tests: Lilliefors, Anderson–Darling and Jarque–Bera with a significance level $\alpha = 0.05$. The three tests indicated that the distributions of residuals could be accepted as normal. Only the obtained residuals using the *nGMSD* metric did not follow normal distribution.

Regarding the independence of residuals, Table 4 shows the computed DW statistics following (7) for the considered metrics. If the value of this statistic was two, the

residuals were completely independent. For a sample size of 52 eyes and a significance level of $\alpha = 0.05$, the critical values of the statistic were approximately $DW_{(L,\alpha)} = 1.49$ and $DW_{(U,\alpha)} = 1.60$. The obtained DW values indicated that no performed fit led to a negative self-correlation in residuals. The correlations of the decimal subjective VA with the $nMSE$, $nPSNR$, and $PSNR-HVS$ metrics, and with the metrics obtained by the linear $nMSE-nPSNR$ and $nMSE-nPSNR-HVS$ combination, gave DW values below $DW_{(L,\alpha)}$. Such values indicated statistically significant evidence at 95% and the error terms were positively self-correlated. In contrast, all the other used measures or combinations of measures gave DW values over $DW_{(L,\alpha)}$, which indicates that residuals did not positively self-correlate, or the test was inconclusive (for $SSIM$ and $nMSE-FSIM$).

Table 4. Durbin–Watson, WI , and \overline{QCD} indices (in red: those metrics or linear combinations that provide positive self-correlations in residuals).

	$nMSE$	$nPSNR$	$SSIM$	$nGMSD$	$MSSSIM$	$FSIM$	$nPSNR-HVS$	$nMSE-nPSNR$	$nMSE-SSIM$	$nMSE-nGMSD$	$nMSE-MSSSIM$	$nMSE-FSIM$	$nMSE-nPSNR-HVS$
DW	1.30	1.28	1.59	1.73	1.61	1.54	1.42	1.28	1.67	1.62	1.67	1.56	1.35
WI	40.7	39.4	44.7	31.1	42.0	47.7	41.7	38.8	47.4	42.4	45.2	50.8	39.3
\overline{QCD}	0.46	0.12	0.15	0.15	0.13	0.06	0.15	0.46	0.14	0.15	0.25	0.07	0.46

The analysis of models’ performance was completed by calculating the WI and \overline{QCD} indices, which are, respectively defined in (8) and (9), and also presented in Table 4. The higher the WI index, the more accurate the model. Conversely, the lower \overline{QCD} , the less spread the results of a metric, which indicates that the metric was relatively independent of the chosen set of images. Thus, its performance was better.

Of the metrics not ruled out by the DW evaluation, the results showed that the $nMSE-FSIM$ metrics combination gave the best results. It can be accepted that residuals were normally distributed, the WI index was higher (higher coefficient of determination), and the index (QCD) had the second lowest value. It also indicated the smallest estimated error for all the metrics used in this paper.

To summarize, the model that best determined subjects’ VA from the image processing metrics based on a logistic function was the $nMSE-FSIM$ combination:

$$VA = 7.478 \left(\frac{1}{2} - \frac{1}{1 + \exp[8.858(Q_L + 3.770)]} \right) - 0.632 Q_L + 1.011, \tag{10}$$

with $Q_L = -0.944nMSE - 6.714FSIM$ and an uncertainty of ± 0.14 .

4. Conclusions

It is often difficult to subjectively determine subjects’ uncorrected VA (their age, lack of cooperation with subjective measurement, etc.). In other cases, it would be desirable to predict this VA prior to any type of ocular intervention. The proposed method allows such a determination to be made objectively by determining the degradation of the eye’s optical system on a set of natural images.

A method is presented to make an objective assessment of subjects’ decimal VA, which can be determined with low uncertainty. The technique is based on using image quality metrics together with the determination of degraded images by the optical part of the HVS. This technique allows both the objective determination and quantification of visual quality. The evaluation of the correlation of VA with the results of different metrics using public domain images revealed that $FSIM$ best performed of all the individual studied metrics. Furthermore, we propose a linear combination of these metrics that provides efficient and effective results, namely $Q_L = -0.944nMSE - 6.714FSIM$.

It is worth noting that a common criticism of this type of mathematical prediction model is that correlation can be confused with chance. We believe that both numerical results and the justification of the hypothesis (image quality metrics can objectively de-

termine subjects' VA) more than suffice to confirm correlation and causality between the studied variables.

It would be interesting to conduct a more extensive study, such as using other grayscale metrics, including the color factor in images or combinations of metrics, to improve the results.

Author Contributions: Writing—original draft preparation, J.E.T. and J.P.R.; Methodology, J.P.R., E.P., C.V.F. and J.E.T.; Software, E.P. and D.M.C.; Formal Analysis, D.M.C., C.V.F., J.E.T. and J.P.R.; Investigation, J.P.R., C.V.F. and E.P.; Resources, E.P. and D.M.C.; Data Curation, J.P.R.; Writing—review and editing, J.E.T. and J.P.R.; Supervision, D.M.C. and E.P.; Project Administration, D.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the University of Alicante for human studies (file number UA-2018-02-19).

Informed Consent Statement: Informed consent was obtained from all the subjects involved in the study.

Data Availability Statement: Data Availability Statements will be available under reasonable request to jorge.perez@ua.es.

Acknowledgments: To the University of Alicante for its support in conducting the measurements.

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

Appendix A

Appendix A.1. The Structural Similarity Index (SSIM)

The SSIM is a metric proposed by Wang and Bovik based on the combination of the luminance and contrast distortion and the loss of correlation between the pixels of the image [8]. The analytical expression of SSIM is shown in expression (A1),

$$SSIM(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (A1)$$

being μ_x and μ_y the mean values of the luminances of the two compared X and Y images, σ_x and σ_y are the standard deviations of the luminances and σ_{xy} is the covariance between the two images. The values C_1 , C_2 and C_3 are constants used to avoid instability when the denominator or denominators approaches zero and are given by:

$$C_1 = (K_1L)^2; \quad C_2 = (K_2L)^2; \quad \text{and} \quad C_3 = C_2/2, \quad (A2)$$

L is the dynamic range of the image ($L = 255$ for 8-bit/ pixel grayscale images), and K_1 and K_2 are two scalar constants with a value less than unity, usually 0.01 and 0.03.

Appendix A.2. The Multi-Scale Structural Similarity Index (MSSSIM)

Wang, Simoncelli and Bovik developed a multiple-scale SSIM, MSSSIM [9] in order to incorporate image details into different image resolutions. In this metric both the reference image and the distorted image are taken as input signals. As in the SSIM, they compare luminance, contrast and correlation between images. The system repeatedly applies a low-pass filter and decreases the resolution of the filtered image by a factor 2. The authors called the original image as Scale 1, and the most distorted image as Scale M (usually $M = 5$), which is obtained after $M-1$ iterations. The comparison between luminance is calculated only on the scale M , $l_M(X, Y)$. However, the contrast and correlation between images are obtained as in SSIM, but each time the image is distorted. Both are denoted by $c_j(X, Y)$ and

$s_j(X, Y)$, respectively, and j is the number of performed iterations. Therefore, MS-SSIM is calculated by combining measurements from different scales using Equation (A3).

$$MS - SSIM(X, Y) = [l_M(X, Y)]^{\alpha_M} \cdot \prod_{j=1}^M [c_j(X, Y)]^{\beta_j} s_j(X, Y)^{\gamma_j}, \tag{A3}$$

where we use the exponents α_M , β_j and γ_j , originally obtained by Wang et al. [9], to adjust the importance of each component.

Appendix A.3. The Gradient Magnitude Similarity Deviation (GMSD)

The gradient is generally calculated by convolving an image with a linear filter, usually a Prewitt filter, along both the horizontal direction, h_h , and the vertical direction, h_v . By performing the convolution of h_v and h_h with the reference image and the distorted image, the horizontal and vertical gradient images of X (reference image) and Y (distorted image) are obtained as follows:

$$\begin{aligned} m_X(i) &= \sqrt{(X * h_v)^2(i) + (X * h_h)^2(i)} \\ m_Y(i) &= \sqrt{(Y * h_v)^2(i) + (Y * h_h)^2(i)} \quad i \in [1, N], \end{aligned} \tag{A4}$$

where i represents any pixel of the image and N is the total number of pixels in the image. With the gradient images, we calculate the gradient magnitude similarity (GMS),

$$GMS(i) = \frac{2m_X(i)m_Y(i) + c}{m_X^2(i) + m_Y^2(i) + c'} \tag{A5}$$

where c is a positive constant that provides numerical stability and controls the contrast response in low gradient areas. In this work, we used a value of $c = 0.0026$, obtained by Xue et al. [11] provided that 8-bit images are considered and with normalized luminance in the range $[0, 1]$. If $m_X(i)$ and $m_Y(i)$ are the same, $GMS(i)$ reaches the maximum value, 1. The GMS map serves as the quality map of the distorted image and reflects the local quality of each small area in the distorted image. The most used way to calculate the quality of a distorted image is to obtain the mean of the elements of the GMS map, $GMSM$,

$$GMSM = \frac{1}{N} \sum_{i=1}^N GMS(i) \tag{A6}$$

Based on the idea that the overall variation of the degradation of local image quality may reflect its overall quality, Xue et al. proposed a metric that evaluates the standard deviation of the GMS map. This metric, gradient magnitude similarity deviation is denoted by $GMSD$,

$$GMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (GMS(i) - GMSM)^2} \tag{A7}$$

The value of this metric is null if there is no distortion. Although the metric does not have higher bounds, we can observe that, for very high values of distortion (normalized difference mean opinion score, $DMOS$, close to unity), the value of this metric is approximately 0.35.

Appendix A.4. The Peak Signal to Noise Ratio Based Human Visual System (PSNR-HVS)

The $PSNR-HVS$ metric is equivalent to the peak signal-to-noise ratio, but it considers the human contrast sensitivity function. It is based on the differences between the coefficients that appear when performing a discrete transform of cosine (DCT) in blocks of 8×8 pixels of both the original and the distorted image. To determine if the differences between DCT coefficients of two images are visually distinguishable, such differences are weighed using a mask based on the quantization table for the JPEG Y color component

obtained considering the contrast sensitivity function. According to [10], the PSNR-HVS metric is more efficient than other metrics and can be expressed as:

$$PSNR - HVS = 10 \log \left(\frac{L}{MSE_H} \right) \tag{A8}$$

where

$$MSE_H = \frac{1}{64(N-7)^2} \sum_{i=1}^{N-7} \sum_{j=1}^{N-7} \sum_{k=1}^8 \sum_{l=1}^8 \left((C^Y[k, l]_{ij} - C^X[k, l]_{ij}) T(k, l) \right)^2 \tag{A9}$$

In the expression (18), $C[k, l]_{ij}$ are the coefficients of the DCT whose upper left coordinates are (i, j) , for both X and Y images. T is a matrix of correction factors proposed in the JPEG algorithm [36,37]. The disadvantage of this metric is that it is not above bounded. It presents zero value if there is no relation between the two compared images and a very high value when they are very similar. When implementing this metric, Ponomarenko assigned a value of 10,000 if the images are the same.

Appendix A.5. The Feature Similarity Index (FSIM)

The last of the used metrics has been the FSIM [12]. It is a full reference metric based on the determination of two characteristics of the compared images: the phase congruence (PC) and the gradient map of the image (GM). The PC is a dimensionless measure of the local structure [38]. It is used as a main feature in FSIM. The theory on PC function offers a simple but biologically plausible model of how mammalian visual systems detect and identify features in an image [39,40]. Considering that the PC is invariant to the contrast but considering that the local contrast of the image affects the perception of the HVS on the quality of the image, the magnitude of the GM is used as a secondary feature to encode contrast information. The PC and GM play complementary roles in the characterization of local image quality. After obtaining the local similarity map, the PC is again used as a weighting function to obtain a unique quality score at each point in the image. According to Zhang et al. [12], the FSIM metric can achieve a correlation with subjective evaluations higher than all IQA metrics do.

The calculation of the FSIM index consists of two stages. In the first stage, the local similarity map is calculated, and then in the second stage, the similarity map is grouped into a single similarity score. Given the X and Y images, the FSIM value for a pixel $\mathbf{x} = (i, j)$ of these images is given by:

$$FSIM_{X,Y}(\mathbf{p}) = \frac{\sum_{\mathbf{p}} [S_{PC}(\mathbf{p}) \cdot S_G(\mathbf{p}) \cdot PC_m(\mathbf{p})]}{\sum_{\mathbf{x}} PC_m(\mathbf{p})} = \frac{\sum_{\mathbf{x}} \left[\left(\frac{2 \cdot PC_X(\mathbf{p}) \cdot PC_Y(\mathbf{p}) + T_1}{PC_X^2(\mathbf{p}) + PC_Y^2(\mathbf{p}) + T_1} \right) \cdot \left(\frac{2 \cdot GM_X(\mathbf{p}) \cdot GM_Y(\mathbf{p}) + T_2}{GM_X^2(\mathbf{p}) + GM_Y^2(\mathbf{p}) + T_2} \right) \cdot PC_m(\mathbf{p}) \right]}{\sum_{\mathbf{x}} PC_m(\mathbf{p})} \tag{A10}$$

where $PC_m(\mathbf{x}) = \max(PC_X(\mathbf{x}), PC_Y(\mathbf{x}))$. T_1 and T_2 are two positive constants used to increase the stability of the $S_{PC}(\mathbf{p})$ and $S_G(\mathbf{p})$ functions [12].

References

1. Guirao, A.; Williams, D.R. A method to predict refractive errors from wave aberration data. *Optom. Vis. Sci.* **2003**, *80*, 36–42. [[CrossRef](#)] [[PubMed](#)]
2. Cheng, X.; Bradley, A.; Thibos, L.N. Predicting subjective judgment of best focus with objective image quality metrics. *J. Vis.* **2004**, *4*, 310–321. [[CrossRef](#)] [[PubMed](#)]
3. Marsack, J.D.; Thibos, L.N.; Applegate, R.A. Metrics of optical quality derived from wave aberrations predict visual performance. *J. Vis.* **2004**, *4*, 322–328.
4. Watson, A.B.; Ahumada, A.J. Predicting visual acuity from wavefront aberrations. *J. Vis.* **2008**, *8*, 17. [[CrossRef](#)] [[PubMed](#)]
5. ITU-T, Recommendation P.800.1; Mean Opinion Score (MOS) Terminology. International Telecommunication Union—Telecommunication Standardization Sector: Geneva, Switzerland, 2016.

6. Wang, Z.; Bovik, A.C. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Process. Mag.* **2009**, *26*, 98–117. [[CrossRef](#)]
7. Ivkovic, G.; Sankar, R. An algorithm for image quality assessment. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '04, Florence, Italy, 4–9 May 2004.
8. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
9. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multi-scale structural similarity for image quality assessment. In Proceedings of the IEEE Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 9–12 November 2003; pp. 1398–1402.
10. Egiiazarian, K.; Astola, J.; Ponomarenko, N.; Lukin, V.; Battisti, F.; Carli, M. New full-reference quality metrics based on HVS. In Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, AZ, USA, 22–24 January 2006; Volume 4, p. 4.
11. Xue, W.; Zhang, L.; Mou, X.; Bovik, A.C. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Trans. Image Process.* **2014**, *23*, 684–695. [[CrossRef](#)]
12. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)]
13. Iskander, D.R. A subjective refraction-based assessment of image quality metric. *Photonics Lett. Pol.* **2011**, *3*, 150–152. [[CrossRef](#)]
14. Pérez, J.; Espinosa, J.; Vázquez, C.; Mas, D. Retinal image quality assessment through a visual similarity index. *J. Mod. Opt.* **2013**, *60*, 544–550. [[CrossRef](#)]
15. Oszust, M. Full-reference image quality assessment with linear combination of genetically selected quality measures. *PLoS ONE.* **2016**, *11*, e0158333. [[CrossRef](#)] [[PubMed](#)]
16. Okarma, K. Combined Full-Reference Image Quality Metric Linearly Correlated with Subjective Assessment. In *Artificial Intelligence and Soft Computing, Proceedings of the 10th International Conference, ICAISC 2010, Zakopane, Poland, 13–17 June 2010*; Rutkowski, L., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6113, pp. 539–546.
17. Ieremeiev, O.; Lukin, V.; Ponomarenko, N.; Egiiazarian, K.; Astola, J. Combined full-reference image visual quality metrics. *Electron. Imaging* **2016**, *14*, 1–10. [[CrossRef](#)]
18. Thibos, L.N.; Hong, X.; Bradley, A.; Cheng, X. Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *J. Opt. Soc. Am. A* **2002**, *19*, 2329–2348. [[CrossRef](#)] [[PubMed](#)]
19. Duane, A. Studies in monocular and binocular accommodation, with their clinical application. *Trans. Am. Ophthalmol. Soc.* **1922**, *20*, 132–157. [[CrossRef](#)] [[PubMed](#)]
20. Sheikh, H.R.; Wang, Z.; Cormack, L.K.; Bovik, A.C. LIVE Image Quality Assessment Database Release 2. Available online: <http://live.ece.utexas.edu/research/quality/subjective.htm> (accessed on 7 March 2023).
21. Le Callet, P.; Atrousseau, F. Subjective Quality Assessment IVC Database. Available online: http://ivc.univ-nantes.fr/en/databases/Subjective_Database/ (accessed on 15 October 2022).
22. Horita, Y.; Shibata, K.; Kawayoke, Y. Toyama-MICT Database. Available online: <http://mict.eng.u-toyama.ac.jp/mict/index2.html> (accessed on 15 October 2022).
23. Spector, R.H. The Pupils. In *Clinical Methods: The History, Physical, and Laboratory Examinations*, 3rd ed.; Walker, H.K., Hall, W.D., Hurst, J.W., Eds.; Butterworths: Boston, MA, USA, 1990; Chapter 58.
24. Goodman, J.W. *Introduction to Fourier Optics*, 3rd ed; W.H. Freeman & Co., Ltd.: New York, NY, USA, 2005.
25. Applegate, R.A.; Lakshminarayanan, V. Parametric representation of Stiles-Crawford functions: Normal variation of peak location and directionality. *J. Opt. Soc. Am. A* **1993**, *10*, 1611–1623. [[CrossRef](#)]
26. Roorda, A. Human visual system-image formation. In *Encyclopedia of Imaging Science and Technology*; Hornak, J.P., Ed.; John Wiley & Sons: New York, NY, USA, 2002; Volume 1, pp. 539–557.
27. Charman, W.N. Wavefront aberration of the eye: A review. *Optom. Vis. Sci.* **1991**, *68*, 574–583. [[CrossRef](#)]
28. Prakash, G.; Srivastava, D.; Suhail, M.; Bacero, R. Assessment of bilateral pupillary centroid characteristics at varying illuminations and post-photopic flash response using an automated pupillometer. *Clin. Exp. Optom.* **2016**, *99*, 535–543. [[CrossRef](#)]
29. Schwiegerling, J. Scaling Zernike expansion coefficients to different pupil sizes. *J. Opt. Soc. Am. A* **2002**, *19*, 1937–1945. [[CrossRef](#)]
30. Dai, G. Scaling Zernike expansion coefficients to smaller pupil sizes: A simpler formula. *J. Opt. Soc. Am. A* **2006**, *23*, 539–543. [[CrossRef](#)]
31. Neter, J.; Wasserman, W.; Kutner, M.H. *Applied Linear Statistical Models*; Irwin: Burr Ridge, IL, USA, 1990.
32. Durbin, J.; Watson, G.S. Testing for serial correlation in least squares regression: I. *Biometrika* **1950**, *37*, 409–428.
33. Durbin, J.; Watson, G.S. Testing for serial correlation in least squares regression. II. *Biometrika* **1951**, *38*, 159–178. [[CrossRef](#)] [[PubMed](#)]
34. Wei, J.; Chen, T.; Liu, G.; Yang, J. Higher-order Multivariable Polynomial Regression to Estimate Human Affective States. *Sci. Rep.* **2016**, *6*, 23384. [[CrossRef](#)] [[PubMed](#)]
35. Bonett, D.G. Confidence interval for a coefficient of quartile variation. *Comput. Stat. Data Anal.* **2006**, *50*, 2953–2957. [[CrossRef](#)]
36. Ponomarenko, N.; Silvestri, F.; Egiiazarian, K.; Carli, M.; Astola, J.; Lukin, V. On between-coefficient contrast masking of DCT basis functions. In Proceedings of the Third International Workshop on Video Processing and Quality Metrics, Scottsdale, AZ, USA, 25–26 January 2007; p. 4.

37. Wallace, G.K. The JPEG still picture compression standard. *IEEE Trans. Consum. Electron.* **1992**, *34*, xviii–xxxiv. [[CrossRef](#)]
38. Kovese, P. Image features from phase congruency. *J. Comp. Vis. Res.* **1999**, *1*, 1–26.
39. Morrone, M.C.; Burr, D.C. Feature detection in human vision: A phase-dependent energy model. *Proc. R. Soc. Lond. B Biol. Sci.* **1988**, *235*, 221–245.
40. Henriksson, L.; Hyvriinen, A.; Vanni, S. Representation of cross-frequency spatial phase relationships in human visual cortex. *J. Neurosci.* **2009**, *29*, 14342–14351. [[CrossRef](#)]

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