

A subjective refraction-based assessment of image quality metrics

D. Robert Iskander*

*Institute of Biomedical Engineering and Instrumentation, Wrocław University of Technology,
Wybrzeże Wyspińskiego 27, 50-370 Wrocław Poland*

Received October 27, 2011; accepted November 08, 2011; published December x, xxxx

Abstract—Tools for quantifying image quality should be benchmarked against subjective human assessment. It is expected that image quality metrics (IQMs) that are based on a human visual system should closely correlate with an observer's perception of vision quality. Performances of a range of popular IQMs were evaluated using subjective and objective measures of optimal refraction from 120 eyes. An optimization algorithm was devised in which objective refraction corresponding to the peak of a given IQM was sought. That optimized objective refraction was then correlated with the corresponding subjective refraction. The results show that although most IQMs result in a similar performance for defocus, those which only use the second order statistical information do not correlate well with a subjective assessment of image quality.

Assessing image quality has been of significant concern in many optical and image processing applications [1, 2]. For example, image compression is an important tool for saving storage space or for increasing the transmission speed in a communication system. This compression must be performed in such a way that no important information is lost and the quality of the original image is well preserved [3]. Unlike one-dimensional signals, images are often viewed by humans. Thus, the results of compression cannot be purely evaluated on the basis of statistical information present in the compressed image but often need to be bench-marked against subjective human assessment [4]. In view of this, significant efforts have been made towards the development of image quality metrics (IQMs) that are based on the human visual system (HVS) [5-8]. However, not all of these metrics have been rigorously bench-marked against subjective human assessment.

Recently, assessing image quality has gained a significant interest in the fields of optometry and vision science. Using retinal IQMs to relate objective measures of visual performance to those subjectively achieved is of particular interest [9-11]. Since the optical system of the eye is assumed to be linear, the retinal image quality assessment can be essentially reduced to the comparison of the diffraction limited point spread function (PSF) and the actual aberrated PSF or their respective Fourier transformations, the Optical Transfer Functions (OTFs).

Interestingly, the two fields of image processing and visual sciences sometime share the same set of techniques for evaluating image quality such as, for example, the use of entropy [8,10,12].

The aim of the study was to assess the applicability of a range of IQMs to retinal images in the context of assessing human vision quality. In particular, evaluation of their applicability in objectively assessing the optical performance of a real human eye was desired. Of interest was also to show which of the popular metrics used in image processing can truly mimic the human visual system and provide some insight into the development of future metrics that may be of use in both fields.

In this study, data from the subjectively and objectively measured refractions from 120 eyes of 60 normal subjects were used (mean age = 23 years). The spherical refractive correction ranged from -8.5 D to $+1.5$ D, while the cylindrical correction ranged from -3.75 D to 0 D. All subjects underwent standard subjective refraction in mesopic luminance conditions. During refraction, the subjects viewed the 6/6 line of a high contrast Bailey-Lovie acuity chart [13] and provided subjective feedback to achieve optimal image quality. The same clinician performed all subjective refractions.

The objective refraction was assessed from the wavefront aberrations measured with a Hartmann-Shack wavefront sensor (COAS, Wavefront Sciences, Inc., Albuquerque). A set of Zernike polynomials up to the 8th radial order (as per ANSI standard [14]) was fitted to each acquired wavefront aberration using the natural pupil size which closely corresponded to that of subjective mesopic refraction.

A number of IQMs from both fields have been considered. They were divided into two groups: those calculated in the spatial domain (based on PSF) and those calculated in the spatial frequency domain (based on the OTF). The first group included

1. Image fidelity (IF)
2. Image entropy
3. Mean structural similarity (MSSIM)
4. Weighted mean structural similarity (WMSSIM)
5. Neural sharpness (NS)

* E-mail: robert.iskander@pwr.wroc.pl

Most of the above metrics are well known in the image processing community. First, the image fidelity [2] was chosen as it is the most basic metric. However, its performance as an image quality indicator has been criticized [7, 15]. Image entropy [8, 10, 12] and structural similarity [6, 7] are becoming the current standard for image quality assessment while neural sharpness is the only one that is exclusively used in vision science [16]. It is, essentially, a Strehl ratio calculated in the PSF domain, where the PSF is weighted by a Gaussian function to mimic the neural function of the visual system. This metric has shown good correlation with subjective assessment of visual quality [11, 16].

The second group included metrics that are currently exclusively used in vision sciences

1. Visual Strehl ratio based on the modulation transfer function (VSMTF)
2. Visual Strehl ratio based on the optical transfer function (VSOTF)
3. Augmented visual Strehl ratio based on the optical transfer function (VSOTFA)

These metrics were included in the performance analysis as they showed, in several studies [9, 11], good correlation with subjective visual performance.

Although it is possible to directly estimate objective refraction from measured wavefront aberrations using the so-called pupil plane measures [17], it is not possible to do so from the subsequent reconstruction of the PSF in the retinal image plane. This is of some concern as it is well accepted in the vision science community that the IQMs calculated in the retinal image plane are more appropriate than those calculated in the pupil plane. In particular, the former can take into account such elements of vision as cone directionality and contrast sensitivity function of the eye.

Thus, in order to objectively calculate the refraction of the eye from the PSF (or OTF) the following optimization algorithm has been devised. In the first step, for each of the measured wavefronts, an ocular refractive power map that represents the distribution of the refractive power across the pupil is calculated [17-19]. Next, the refractive power was used to estimate the best spherocylinder [20], which was then transformed to a spherocylindrical wavefront. In the following, the resulting wavefront was subtracted from the original wavefront producing a wavefront with a zero spherocylinder. Finally, numerous spherocylindrical combinations ranging from -1 D to +1 D in 0.25 D steps of spherical power, -1 D to 0 D in 0.25 D steps of cylinder power and axes between 0 and 175 degrees in 5-degree steps were added to the derived residual wavefront. The PSF and the corresponding OTF were then calculated for each of the spherocylindrical combinations from which the given IQM was estimated. A spherocylindrical combination that resulted in the peak of

the IQM (minimum for image entropy and maximum for all the other metrics) with the addition of the originally estimated spherocylinder was then used for comparison with subjective refraction.

It is important to note that in order to compare two spherocylinders (sphere, cylinder, and axis) one needs to transform each of them into three orthogonal refractive power vector components M, J0 and J45, where M denotes the mean spherical power while J0 and J45 describe astigmatic components at 0 and 45 degrees, respectively [21]. Since measurement errors are present both in subjective and objective assessments of vision quality, orthogonal linear regression was used to establish correlation between subjectively assessed power vector components and their objectively assessed counterparts. For each of them the slope, intercept, and Pearson's correlation coefficient r^2 were calculated.

The results of correlation between subjective and objective refractions in terms of refractive power vector components, M, J0 and J45 are given in Table 1. Please note that image entropy (indicated in bold font) has achieved the best global correlation for all refractive power vector components.

Table 1. Correlation between subjective and objective refractions in terms of refractive power vectors for a range of IQMs.

METRIC	M			J ₀			J ₄₅		
	slope	intercept	r^2	slope	intercept	r^2	slope	intercept	r^2
Maloney's	1.05	-0.36	0.97	1.14	0.04	0.78	1.04	0.00	0.58
IF	1.04	-0.35	0.97	1.33	0.04	0.64	1.44	0.01	0.20
Entropy	1.03	-0.37	0.97	1.14	0.04	0.82	1.20	0.00	0.61
MSSIM	1.03	-0.62	0.95	1.43	0.03	0.58	1.28	0.02	0.05
WMSSIM	1.04	-0.57	0.95	1.42	0.03	0.61	1.23	0.01	0.04
NS	1.03	-0.42	0.97	1.15	0.04	0.82	1.06	0.01	0.59
VSMTF	1.03	-0.31	0.97	1.19	0.04	0.78	1.30	0.00	0.53
VSOTF	1.04	-0.32	0.97	1.22	0.06	0.73	1.20	0.01	0.51
VSOTFA	1.03	-0.34	0.97	1.20	0.04	0.78	1.18	0.00	0.58

*The Maloney's spherocylinder estimate is not an IQM and it is derived directly from the objectively measured wavefront aberration.

All considered IQMs have achieved high correlations between the subjective and objective measures of vision quality for the mean spherical refractive power vector component ($r^2 > 0.95$). All of them also showed a significant level of intercept of up to -0.62 D indicating the so-called "instrument myopia" [22]. This is a common problem in wavefront sensors that do not have an open field-view in which the subject's level of accommodation (focusing) is not fully controlled.

All of the IQMs that are used in vision sciences (i.e., image entropy, NS, VSMTF, VSOTF, and VSOTFA) showed a relatively high correlation for the astigmatic refractive power vector components J0 and J45, but were not significantly different to that already achieved by the Maloney's estimate of the spherocylinder. It should be noted that this estimate of the spherocylinder is not an IQM and it is derived directly from objectively measured

wavefront aberration [17], making the whole process of objective refraction computationally less expensive. This suggests that there is little need for the vision science community to develop more complex IQMs. Those that are popular in the image processing community (i.e., IF, MSSIM, and WMSSIM) have resulted in much lower correlation between the subjective and objective measures of astigmatism than that achieved for the vision science metrics. This was a surprising result for the structural similarity metrics. In a preliminary study, it was found they were able to well differentiate between varying levels of astigmatism and its axes when applied to digital images with artificially induced astigmatism. Further investigation showed that the window size in the MSSIM as well as the shape of the weighting function in the WMSSIM had little or no effect on correlation results. It was somehow an unexpected result.

Defocus and astigmatism are the two major aberrations found in human eyes. They correspond to the second order information in the expansion of a wavefront into a series of Zernike polynomials and are equivalent to second order statistics (the spread) of the PSF. The existence of higher order aberrations, such as coma and spherical aberration, in the human eyes is well documented [23, 24]. These normally uncorrected higher order aberrations are evident even in normal healthy eyes. They reduce the quality of retinal image but increase the depth of field by about 0.25 D and reduce the required precision of accommodation [25]. There is a complex relationship between lower order aberrations, higher order aberrations, accommodation response, and retinal image quality [26].

Since the structural similarity metrics are based on first and second order statistical information, it was of interested to see how they would perform if one has measured only the defocus and astigmatism components of aberration. To do so, a set of Zernike polynomials of 2nd radial order was fitted to each acquired wavefront aberration and the optimization algorithm described above was run for the MSSIM, WMSSIM, and the previous best metric – the image entropy. The results showed conclusively that all three IQMs arrived at the same poor correlation between the subjective and objective measurements of refraction. This means that in the presence of higher order aberrations, IQMs that are only based on first and second order statistical information will not adequately mimic the subjective human assessment of image quality.

The study indicated that image entropy, an IQM that considered all statistical information, shows the best correlation with the subjective assessment of visual quality among the considered IQMs. However, it can be speculated that since the amounts of higher order aberrations in the normal human eye decrease with the order [27], the inclusion of up to fourth order statistical information into an IQM should be sufficient for it to adequately mimic the optical performance of a normal healthy eye.

References

- [1] A.J. Ahumada, Jr., SID International Symposium Digest of Technical Papers **24**, 305 (1993).
- [2] A.M. Eskicioglu, P.S. Fisher, IEEE Trans. Commun. **43**, 2959 (1995).
- [3] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, IEEE Trans. Image. Proc. **1**, 205 (1992).
- [4] Z. Wang, A.C. Bovik, L.Lu, Proc. IEEE ICASSP, 3313 (2002).
- [5] D.J. Granrath, Proc. IEEE **69**, 552 (1981).
- [6] Z. Wang, A.C. Bovik, IEEE Signal Process. Lett. **9**, 81 (2002).
- [7] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, IEEE Trans. Image Proc. **13**, 600 (2004).
- [8] H.R. Sheikh, A.C. Bovik, IEEE Trans Image Process. **15**, 430 (2006).
- [9] X. Cheng, A. Bradley, L.N. Thibos, J. Vis. **4**, 310 (2004).
- [10] L. Chen, B. Singer, A. Guirao, J. Porter, D.R. Williams, Optom. Vis. Sci. **82**, 358 (2005).
- [11] J.D. Marsack, L.N. Thibos, R.A. Applegate, J. Vis. **4**, 322 (2004).
- [12] A. Guirao, D.R. Williams, Optom. Vis. Sci. **80**, 36 (2003).
- [13] I.L. Bailey, J.E. Lovie, Am. J. Optom. Physiol. Opt. **57**, 378 (1980).
- [14] *American National Standards for Ophthalmics. Methods for reporting optical aberrations of eyes.* ANSI Z80.28 (2004).
- [15] D.A. Silverstein, J.E. Farrell, Proc. IEEE ICASSP, 881 (1996).
- [16] L.N. Thibos, X. Hong, A. Bradley, R.A. Applegate, J. Vis. **4**, 329 (2004).
- [17] D.R. Iskander, M.J. Collins, B. Davis, R. Franklin, Ophthal. Physiol. Opt. **27**, 245 (2007).
- [18] J. Nam, L.N. Thibos, D.R. Iskander, J. Opt. Soc. Am. A Opt. Image. Sci. Vis. **26**, 1035 (2009).
- [19] J. Nam, L.N. Thibos, D.R. Iskander, Clin. Exp. Optom. **92**, 194 (2009).
- [20] R.K. Maloney, S.J. Bogan, G.O. 3rd Waring, Am. J. Ophthalmol. **115**, 31 (1993).
- [21] O. T. Salmon, R.W. West, W. Gasser, T. Kenmore, Optom. Vis. Sci. **80**, 6 (2003).
- [22] R.T. Hennessy, J. Opt. Soc. Am. **65**, 1114 (1975).
- [23] W.N. Charman, Optom Vis Sci. **68**, 574 (1991).
- [24] D.A. Atchison, Clin Exp Optom. **87**, 138 (2004).
- [25] M.J. Collins, T. Buehren, D.R. Iskander, Vis. Res. **46**, 196 (2006).
- [26] T. Buehren, M.J. Collins, Vis. Res. **46**, 1633 (2006).
- [27] A. Guirao, J. Porter, D.R. Williams, I.G. Cox, J. Opt. Soc. Am. A Opt. Image. Sci. Vis. **19**, 620 (2002).