Assessment of Image Quality Using a Pseudophakic Eye Model for Refractive Evaluation

Filomena Ribeiro[1,2,3], Antonio Castanheira-Dinis[2], João Miguel R. Sanches[4], and João M. Dias[1]

[1] GoLP/Instituto de Plasmas e Fusão Nuclear-Laboratório Associado, Instituto Superior Técnico, Technical University of Lisbon, Portugal
[2] Visual Sciences Research Centre, University of Lisbon, Lisbon, Portugal
[3] Hospital da Luz, Lisbon, Portugal
[4] Institute for Systems and Robotics, Department of Bioengineering, Instituto Superior Técnico, Technical University of Lisbon, Portugal
filomenajribeiro@gmail.com

Abstract. In the refractive assessment by optical evaluation based on ray-tracing, the definition of the best focus plane remains a challenge. We simulated 100 pseudophakic eye models using a Montecarlo analysis with ray-tracing evaluation. The image quality resulting from optimization with the Visual Strehl ratio computed in frequency domain weighted by the neural contrast sensitivity function (VSMTF), a metric that has been shown to have a good correlation with defocus detection by the human eye, and with the Root-Mean-Square of Wavefront (RMSW) error, the most commonly used optimization metric, was assessed. For objective assessment, we designed an index to detect increasing stages of defocus. For subjective assessment, we designed a force choice test that was completed by 20 observers. Results show that both for objective and subjective evaluation, VSMTF performed better than RMSW. Therefore, VSMTF is a good metric for the refractive assessment of human eye models with ray-tracing.

Keywords: Defocus, eye models, ray tracing, visual optics, optical aberrations, metrics of optical quality.

1 Introduction

Vision quality has gained increasing importance in assessing the results from both refractive and cataract surgery. However, the assessment of vision quality is a challenge, since it cannot be measured directly. Optical properties are defined by anatomical characteristics of the eye, thus determining the quality of the image formed at the retina and setting limits for functional tasks, such as resolution and contrast detection. Finally, the image is processed by the Central Nervous System, establishing the final perception of the initial visual stimulus. However, the human eye has optical aberrations that limit the quality of image. Therefore, the question remains: when considering all the optical aberrations of the human eye, how does one define the best focus plane when correcting for defocus. The most commonly used metric on the evaluation

J.M. Sanches, L. Micó, and J.S. Cardoso (Eds.): IbPRIA 2013, LNCS 7887, pp. 543–550, 2013.

of optical systems with ray-tracing, the Root-Mean-Square of Wavefront (RMSW) error, is highly effective in describing optical systems close to the limit, by diffraction, but not in more imperfect systems such as the human eye [1]. The Visual Strehl ratio computed in frequency domain weighted by the neural contrast sensitivity function (VSMTF) has been shown to be one of the best metrics to estimate defocus [2]. This metric is optimized by the same level of defocus that gives the best visual acuity and successfully predicts the best subjective focus plane. Moreover, metrics based on the Modulation Transfer Function (MTF) optimization are effective for both poorly and well-corrected systems [3].

In a previous paper we have reported the use of this metric on the evaluation of pseudophakic models with ray-tracing, and on IntraOcular Lens (IOL) power estimation when only defocus is corrected [4].

The main goal of this paper is validating this methodology concerning refractive assessment by ray-tracing. In order to do so, we propose two strategies, an objective and a subjective .method.

2 Methods

We have conducted a two stage analysis, schematically depicted in Figure 1. The designed image quality objective evaluation index and subjective assessment were applied to the images generated by 100 Montecarlo pseudophakic eye models. Eye model definitions were based on the Liou-Brennan eye model [5], and the several parameters were set as previously described [4].

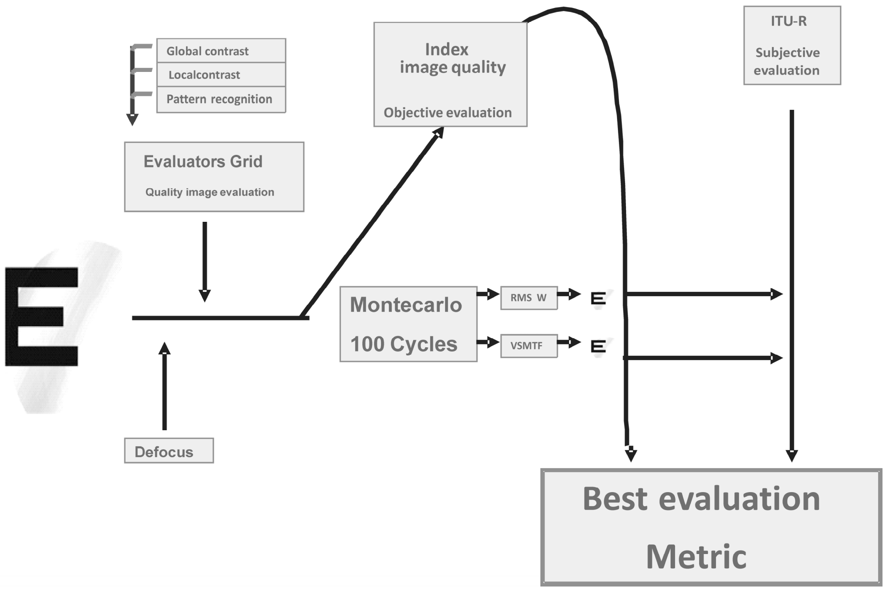

Fig. 1. Schematic representation of the several steps of data analysis performed

2.1 First Stage

Design of the Objective Evaluation Grid

An image evaluation grid was designed, in Matlab®, considering a) Intensity and contrast metrics, global and local contrast and b) Pattern recognition and correlation with object image. Table 1 describes all the evaluators used to design this image evaluation grid.

Table 1. Objective image evaluation grid designed in Matlab®

Objective Image Quality Evaluator	
Intensity and contrast evaluators	
1	Histogram variance
2	Intensity standard deviation
3	Difference between intensity and mean intensity divided by standard deviation
4	Difference between maximum and minimum intensity divided by mean intensity
5	Local difference between maximum and minimum intensity divided by mean intensity
6	Weber contrast
7	Michelson contrast
8	Difference between intensity and local standard deviation divided by local standard deviation
9	Total Variation-Sum of the components of the two-dimensional numerical gradient
10	Local standard deviation of image
11	Local contrast prior-Relation between local gradient and local contrast
12	Local contrast of gray-level co-occurrence matrix
Pattern recognition and correlation with object image	
13	Cross-correlation between image and object
14	Canny method to find edges by looking for local maxima of the gradient of Image
15	Mutual information between object and image canny edges
16	Difference phase angle between object and image

Intensity and contrast metrics, global and local contrast.

Given our aim was to simulate the visual task that constitutes Snellen's visual acuity, we have taken into account the model proposed by Atchison et al [6] to describe the criteria used by the human visual system to perceive the blur of a letter depending on the size of that letter, local contrast and the detection of change in edge gradient, in addition to global contrast. Therefore, we have included measures of contrast applied to small image windows, for local assessment, as well as measures to detect edges and assess gradients.

Moreover, in the case of small Snellen's optotypes of variable contrast over a homogeneous background, the Weber contrast is recommended [7]. Also, in cases of dark and bright alternating equivalent patterns, as is the letter "E" used in the present study, the Michelson's measurement is recommended. Therefore, these two evaluators were included in our grid [7].

Pattern recognition and correlation with object image.
Visual performance evaluation, given by Snellen's visual acuity, implies the modulation of the whole visual process, including pattern recognition. For more complex patterns, such as optotype letters, detection can also be affected by the Optical Transfer Function (OTF) phase, which, added to contrast loss, sharply decreases recognition ability [8]. Therefore, we applied contour detection algorithms (Canny method) [9] to the image, and evaluated the correlation between object and image obtained, using matrix correlation and mutual information [10]. Evaluation of OTF's phase difference was performed by previous application of a Fourier transform.

Assessment of the Evaluation Grid Behavior for Progressive Stages of Defocus
In order to establish an index adapted to defocus assessment, we have generated 23 progressive stages of defocus, simulated through the application of a rotationally symmetric Gaussian low pass filter, with standard deviation sigma (positive). This was applied to the letter "E" and each obtained image was assessed by the previously defined grid.

Establishment of the Assessment Index
The evaluation index is an estimation of the defocus, d, that is here modeled as a linear combination of the measures, $F = \{f_k\}$, listed on Table 1,

$$d = \sum_{k=1}^{16} \alpha_k f_k = \mathbf{a}^T F \tag{1}$$

where the optimum coefficients $\mathbf{a} = \{\alpha_k\}$ are estimated by the least square (LS) methods,

$$a^* = \arg \min_{\mathbf{a}} \sum_i \left(F_i^T \mathbf{a} - d_i\right)^2$$

with d_i the i^{th} defocus value tested, from a series of increasing synthetic defocused images, and F_i the vector with the corresponding measures (see Table 1). The optimum vector of coefficients is given by $a^* = \mathbf{\Phi}^+ \mathbf{d}$ where $\mathbf{\Phi}^+$ is the pseudoinverse Moore-Penrose matrix and $\mathbf{d} = \{d_i\}$ is the vector of defocus values.

This assessment index thus characterizes the image quality that maximizes the visual ability, measured by Snellen's chart, and is capable of detecting different stages of defocus, through the estimated vector $\mathbf{a}^*$.

2.2 Second Stage

Establishment of the 100 Cases for Objective and Subjective Assessment

We have simulated 100 pseudophakic eyes using the Montecarlo method, considering all parameters, and letting them vary randomly and simultaneously within their physiological range, following a Gaussian distribution. For each model, we have optimized the IOL radius considering two different metrics for optimization, VSMTF and RMSW, thus obtaining the images to be assessed.

Image Quality Evaluation

For each Montecarlo case, the obtained OTF was used to compute each image appearance of the letter "E". Each image was evaluated according to:

1. the objective assessment index established on the first stage;
2. the subjective assessment of 20 observers, following the Recommendation ITU-R BT.500-11 [11], which defines the general and specific viewing conditions for subjective assessments. Observers had to choose between the two images resulting from VSMTF and RMSW optimizations for each eye model, according to item 6.2.4.3 of Recommendation ITU-R BT.500-11, establishing the one that was less blurred and in which the pattern of the letter "E" was more clearly identified, having the option of answering "Don't know".

2.3 Used Optimization Metrics

For each of the 100 Montecarlo cases two metrics were used to optimize IOL dioptric power:

- RMSW: the most commonly used metric to evaluate optical systems, in which the merit function minimizes the optical path difference with respect to the shifted and tilted reference sphere that minimizes the RMS wavefront error;
- VSMTF: a Figure Of Merit (FOM) defined in order to minimize the difference of MTF values with respect to the diffraction limit values, attributing different weights to different frequencies [4].

2.4 Data and Statistical Analysis

For the objective evaluation, assessment index values were summed for each image of the simulated 100 pseudophakic eyes and then the average and standard deviation of this sum was taken. The number of cases that objectively performed better for each metric was also accounted for. For the subjective evaluation, the number of times each image was chosen was analysed individually for each image and each observer. Total scores obtained from each of the observers were also compared.

Means were compared using t-tests. The number of cases that objectively performed better for each metric were compared using the Mann-Whitney U test. Correlations between objective and subjective assessment were evaluated using Spearman's ρ. Tests were considered significant at $p < 0.05$ significance level (two-tailed).

3 Results

In this section the evaluation of the objective and subjective assessment methods are described with synthetic and real data.

3.1 Objective Evaluation of Image Quality

The average of the summed index values for each of the simulated 100 pseudophakic eyes considering VSMTF and RMSW shows that VSMTF performed better than RMSW ($p = 0.032$). Analysis of individual cases showed that VSMTF was better in 77 cases whilst RMSW was better in 23 cases.

3.2 Subjective Evaluation of Image Quality

We used a small population sample of 20 healthy individuals, who subjectively assessed the two images resulting from RMSW and VSMTF optimizations, according to item 6.2.4.3 of Recommendation ITU-R BT.500-11 [11]. VSMTF performed better in 90.2% of cases and RMSW in 2.9%. The answer "Don't know" was chosen in 7.0% of cases. VSMTF also showed the highest absolute concordance between observers – Figure 2.

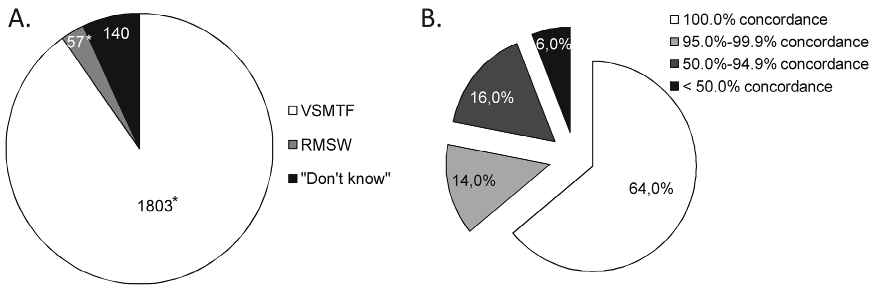

Fig. 2. A. Absolute number for each alternative of subjective choice. *p<0.001. **B.** Percentage of cases for which the total score showed or not absolute concordance between observers regarding VSMTF choice.

3.3 Correlation between Objective and Subjective Assessment of Image Quality

There was a significant correlation between the number of times the VSMTF image was chosen by the observers and objective RMSW-VSMTF index ($\rho = 0.491$, $p<0.001$, $n=100$). Also, the number of times the answer "Don't know" was chosen correlated with RMSW-VSMTF index ($\rho = -0.460$, $p<0.001$, $n=100$). Figure 3 shows the relationship between objective and subjective evaluations.

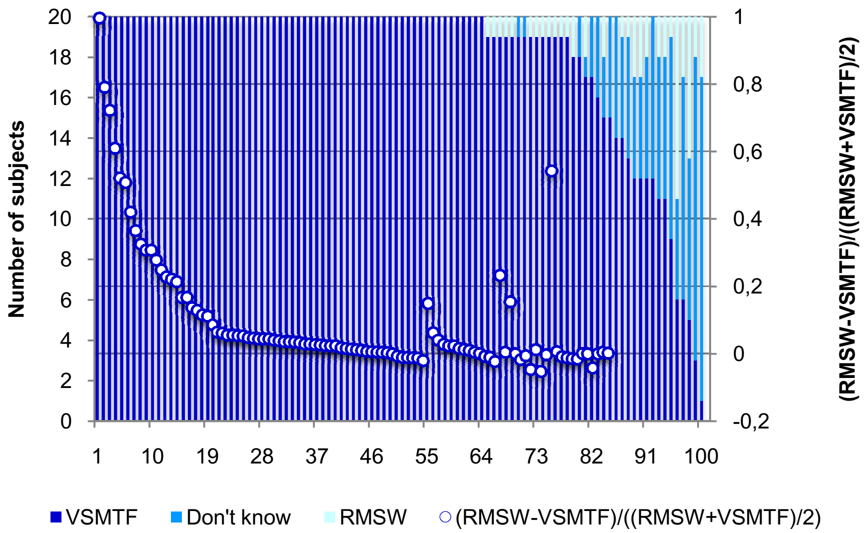

Fig. 3. Objective and subjective assessment of the evaluated images (100 RMSW and 100 VSMTF). Number of times each image was subjectively chosen (left yy axis) and respective objective index (right yy axis) for each of the 100 Montecarlo cases (xx axis).

4 Discussion

The aim of optimization metrics is achieving the best correction for defocus, and the best metrics will be the one that correlates better both with retinal image quality, simulated by OTF convolution with the object image, and with subjective defocus assessment.

The proposed evaluation grid, which is an approximation to the visual task that represents Snellen's visual acuity, and the index resulting from it, allows the assessment of the best focus plane and focus degree, corresponding to the intended evaluation of the image quality obtained in each optimization. Results show that, when considering the objective index, 77 cases had a better image quality with VSMTF, being statistically different from the 23 cases obtained with RMSW.

Given vision is a complex phenomenon and thus the definition of best image quality is not sufficient, it is necessary that this theoretical definition corresponds to the subjective assessment of the human eye. Subjective assessment, for which 90.2% of the cases were better with VSMTF (Figure 2A), not only confirms results obtained for the objective index but also correlates with objective assessment. Correlation analysis (Figure 3) shows that a better objective VSMTF index corresponds to a better perceived image. Moreover, the worse the objective VSMTF index or the smaller the difference between RMSW and VSMTF indices, the more the subject evaluating the image chooses the answer "don't know". Therefore, the proposed VSMTF metric results in a better image than the RMSW metric, both objectively and subjectively.

These results have prompted us to further continue this line of research, and we intend, in the future, to improve the objective assessment index and the chosen evaluators, in order to attain a better correlation with subjective assessment.

5 Conclusions

Visual perception is highly subjective and involves many aspects of image quality, and therefore improved visual quality metrics are needed and must incorporate neural factors and subjective perception. Results show that using VSMTF for the assessment of optical models of the human eye with ray-tracing is better than RMSW, both concerning objective and subjective image assessment.

The human eye has relatively low optical quality when compared to simulated optical systems, and therefore standard metrics of optical quality, such as RMS wavefront error, may not be the most useful metrics for predicting the quality of vision or the optical limits to visual performance. New metrics of optical/neural performance, that correlate better with clinical measures of visual performance, such as VSMTF, need to be adopted in the refractive assessment by ray-tracing.

References

[1] Mouroulis, P.: Aberration and image quality representation for visual optical systems. In: Mouroulis, P. (ed.) Visual Instrumentation: Optical Design and Engineering Principle, pp. 27–68. McGraw-Hill, New York (1999)

[2] Cheng, X., Bradley, A., Thibos, L.N.: Predicting subjective judgment of best focus with objective image quality metrics. J. Vis. 4(4), 310–321 (2004)

[3] Dobson, S.J., Cox, A.: Fast image-quality-based optimization of optical systems. Appl. Opt. 37(34), 8008–8011 (1998)

[4] Ribeiro, F.J., Castanheira-Dinis, A., Dias, J.M.: Personalized pseudophakic model for refractive assessment. PloS One 7(10), e46780 (2012)

[5] Liou, H.L., Brennan, N.A.: Anatomically accurate, finite model eye for optical modeling. J. Opt. Soc. Am. A Opt. Image Sci. Vis. 14(8), 1684–1695 (1997)

[6] Atchison, D.A., Smith, G.: Optics of the Human Eye. Butterworth-Heinemann (2000)

[7] Zanlonghi, X.: Sensibilite au Contrast: Etude Comparative des Appareillages Actuels. Coup d'Oeil 32, 70–74 (1991)

[8] Nestares, O., Navarro, R., Antona, B.: Bayesian model of Snellen visual acuity. J. Opt. Soc. Am. A Opt. Image Sci. Vis. 20(7), 1371–1381 (2003)

[9] Canny, J.: A Computational Approach To Edge Detection. IEEE Trans. Pattern Analysis and Machine Intelligence 8(6), 679–698 (1986)

[10] Korn, G.A., Korn, T.M.: Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review, pp. 613–614. Dover Publications, New York (2000)

[11] RECOMMENDATION ITU-R BT.500-11. Methodology for the subjective assessment of the quality of television pictures (2011), http://www.dii.unisi.it/~menegaz/DoctoralSchool2004/papers/ITU-R_BT.500-11.pdf