A NEW DECOMPOSITION BASIS FOR THE CLASSIFICATION OF ABERRATIONS OF THE HUMAN EYE

LAURENT DUMAS\textsuperscript{1}, DAMIEN GATINEL\textsuperscript{2} AND JACQUES MALET\textsuperscript{3}

Abstract. The optical defects of the human eye are analyzed and classified by studying the corresponding ocular wavefront. After presenting the orthogonal basis, called the Zernike basis, which is currently used for the wavefront analysis, a new mathematical decomposition basis is built. In this new basis, the phase errors due to the low degree aberrations terms are totally separated from the high order terms, in order to improve the relevance of the clinical interpretation through a better appreciation of the low vs high aberration wavefront components. Various clinical examples (myopia, keratoconus) clearly show the interest of this new basis compared to a diagnosis based on the Zernike decomposition.

Résumé. Les défaut optiques de l’œil humain sont analysés et classés à partir des caractéristiques mathématiques du front d’onde de l’œil considéré. Après avoir présenté la méthode actuelle basée sur la décomposition du front d’onde dans la base orthonormale de Zernike ainsi que certaines de ses limitations, on propose ici une nouvelle base de décomposition basée sur une séparation stricte des effets de bas degré et de haut degré. Différents cas cliniques présentés dans la dernière section permettent de mettre en évidence l’intérêt de cette nouvelle base de décomposition.

1. VISION AND OPTICAL ABERRATIONS

1.1. The visual process

Vision is a complex process that can be grossly divided into two steps: an optical step and then a neural one. After being refracted by the cornea, part of the light which enters into the eye passes through the pupil, whose diameter is variable. This light then passes through the crystalline lens, acting as a focusing element with a curvature that can be modified by the action of ciliary muscles in order to form a sharp image on the retina of objects located at various distances from the eye (accommodation principle). The visual information is then transmitted to the brain via the optic nerve.

In the eye, the light travels through various media with different index values. For a monochromatic light wave, its wavelength is decreased in denser media (i.e. with higher refractive indices) whereas its frequency is unchanged. The optical path length is defined as the number of wavelengths between the light source and the image plane. It is equal to the product of the geometric length of the path which light follows, and the index of refraction of the medium through which it propagates. If this distance is the same for all the light rays emitted by a single point source, the quality of the obtained image is optimal, as the latter is created by focusing light waves with a maximal constructive interference at the plane of observation.

\textsuperscript{1} Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, 78035 Versailles, France.
\textsuperscript{2} Department of Anterior Segment and Refractive Surgery, Fondation Rothschild, 75019 Paris, France.
\textsuperscript{3} PhD, no affiliation.
A wavefront is a surface over which an optical wave has a constant phase. In 3D space, a point source emits wavefronts that are spherical in shape, which propagate radially outward. The center of these wavefronts is the point source. The optical rays are materializing the direction of propagation of the wavefront, and are always perpendicular to the surface of the wavefront at each point [5].

The example of a star in a night sky, can be taken as an elementary light source. Only a small portion of the light waves emitted from the star can be captured by an eye gazing at it. Although the wavefronts produced by the star considered as an elementary point are always concentric spheres in principle, the radii of the spheres are so large that they look like plane waves to a terrestrial observer. It is equivalent to consider that the rays that are striking the eye are all parallel. The successive refraction by the cornea and the crystalline lens make these rays ideally focused to the foveal plane into a single location. Such eye is said to be emmetropic.

Hence, neglecting the effect of diffraction, the ideal retinal image of a light point source is a point. The property of an optical system in which rays of light from a single point converge in a single focal point is called "stigmatism". Conversely, if not all these rays are focused in a single point, the image is blurred. In visual optics, this can happen either when the eyeball is too 'long' (myopia) or too 'short' (hyperopia), with respect to its focusing length (see figure 1).

![Figure 1. Eye vision of a point source (a star for instance).](image)

The point spread function (PSF) describes the response of an imaging system to a point source or point object. The degree of spreading (blurring) of the point object is a measure for the quality of an imaging system. For instance, for a myopic eye, the distance to the retina plane is too long for rays emanating from a source located at infinity. The optical path length increases for rays located further away from the optical axis. The PSF is no longer a point but a blur disc. However, the retinal image formed by a myopic eye can be sharp if the point source is located nearer to the eye. In such situation, the peripheral rays have a longer geometrical path to the fovea than the central rays, which thus 'equalizes' the optical path lengths along every refracted ray.

The optical correction for a myopic eye can be achieved with spectacles. The power of the spectacle lens is mainly governed by the focusing error of the rays that enter through the central (paraxial) area of the pupil. Ideally, the corrective glass ensures that all the rays emitted from a distant point source will be focused on the retinal plane. However, due to ocular imperfections, not all rays can be focused on the retina. The quality of vision (e.g. contrast sensitivity) of a given eye is partly governed by the focal plane of rays traveling through the periphery vs the center of the pupil.
1.2. Measuring the wavefront error

Figure 2 gives an example of a retina image made from monochromatic and parallel rays for two different eyes. On the top of this figure, the eye refractive elements (represented schematically by a single lens) are perfect in the sense that after refraction by the cornea and the crystalline lens, the envelope of the wavefronts are spherical and the PSF is analogous to a single point (again neglecting the effect of diffraction). On the bottom of the same figure, the eye has optical imperfections, also called high-order aberrations: some peripheral rays are not focalised in the same plane as the paracentral ones. This corresponds to the presence of phase delays or advances with regards to the reference spherical envelope of the wavefront of a "perfect" eye. As a consequence, the PSF is no longer a single point and the retinal image quality degrades. To quantify such aberrations, the wavefront error function is defined as the representation of the phase differences between a "perfect" wavefront, and the measured wavefront. It is defined on the pupil disc, within which the phase delays and advances are computed.

Figure 3. The wavefront sensing of eyes with or without aberration.
The experimental measurements of ocular wavefronts which are performed in ophtalmology are based on the collection and sampling after reflection on the central retina of a monochromatic light beam (see figure 3) [2,3].

A wavefront reflecting off the retina becomes deformed as it passes through the eye’s optics, crystalline lens and cornea. The aberrated wavefront emerging from the eye contains information about the optical defects it possesses. Imaging techniques allow to quantify the wavefront error of the measured eye. As an example, the wavefront that is collected outside of the eye after foveal reflection of a myopic eye will possess a spherical geometry, which can be modelled by a quadratic function (see an example of such a wavefront on figure 4).

Figure 4. Example of a collected wavefront for an eye with myopia.

The central shape of the collected wavefront thus allows to predict the refractive correction. The power of the spectacle lenses, corresponding to low-order aberrations, can be predicted with acceptable precision in most clinical situation from the osculating curvature of the wavefront, which is expressed as a quadratic function. The presence of peripheral phase errors, expressed as high-order aberrations, will affect the optical quality of the eye. In low light conditions, the pupil dilates, and the wavefront error of human eyes tends to increase dramatically.

Our understanding of the optical quality of the eye has become more accurate with the ability to precisely measure the lower and higher order wave aberrations using ocular wavefront sensing techniques. However, a deeper mathematical analysis of the wavefront function may be needed to better classify and quantify the origin of the corresponding ocular wavefront aberration.

2. The Mathematical Treatment of Wavefronts

On a mathematical viewpoint, the eye wavefront can be seen as a real valued function defined on the \( \mathbb{R}^2 \) unit disc, corresponding to the normalized pupil (assumed to be perfectly spherical, see [6]):

\[
D = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}
\]

The scalar product between two wavefronts of \( \mathcal{C}(D, \mathbb{R}) \) is defined with the expression:

\[
<f, g> = \frac{1}{\pi} \int_D f(x, y)g(x, y)dxdy
\]

In particular, it allows to define the distance from a wavefront to its average, called RMS (Root Mean Square) error:

\[
RMS = ||f - p_0(f)||_D = \frac{1}{\pi} \sqrt{\int_D (f(x, y) - p_0(f))^2dxdy}
\]

where \( p_0(f) = \frac{1}{\pi} \int_D f(x, y)dxdy \).
Denote $\mathcal{P}_D$ the set of polynomial functions on $D$ and $\mathcal{P}_D^{(N)}$ the subset of $\mathcal{P}_D$ made of polynomials of total degree lesser than $N$. A general wavefront can be approximated by its orthogonal projection $P(f)$ on $\mathcal{P}_D^{(N)}$. The degree (respectively valuation) of $P(f)$ is then called degree (respectively valuation) of the projection of $f$.

2.1. The Zernike decomposition basis

The polar coordinates $(r, t) \in [0,1] \times [0,2\pi]$ such that $x = r \cos(t)$ et $y = r \sin(t)$ are more suited to the geometry of the problem and lead to express the wavefront projection as a trigonometric polynomial. The Gram-Schmidt orthonormalisation principle of the radial part of the canonical basis of trigonometric polynomials lead to the definition of the so-called Zernike polynomials (see the reference paper of Zernike [7]). This family is sorted by radial degree, and for a given degree in increasing order of azimuthal frequency. More precisely, the Zernike polynomial $Z_n^m$ is given for any integer $n$ and any $m \in \{-n, -n+2, \ldots, n-2, n\}$ by the expression:

$$Z_n^m(r, t) = R_n^{|m|}(r)T_m(t)$$

where $T_m(t) = \cos(mt)$ if $m \geq 0$ and $T_m(t) = \sin(|m|t)$ if $m < 0$.

After formal computations, it gives:

$$Z_n^m = \begin{cases} \sqrt{2(n+1)} \sum_{0 \leq k \leq \frac{n+m}{2}} (-1)^k \frac{(n-k)!}{k! \left( \frac{n-m}{2} \right)! \left( \frac{n+m}{2} - k \right)!} r^{n-2k} \cos(mt) & \text{if } m > 0 \\ \sqrt{n+1} \sum_{0 \leq k \leq \frac{n}{2}} (-1)^k \frac{(n-k)!}{k! \left( \frac{n}{2} - k \right)! \left( \frac{n}{2} - k \right)!} r^{n-2k} & \text{if } m = 0 \\ \sqrt{2(n+1)} \sum_{0 \leq k \leq \frac{n+m}{2}} (-1)^k \frac{(n-k)!}{k! \left( \frac{n-m}{2} \right)! \left( \frac{n+m}{2} - k \right)!} r^{n-2k} \sin(|m|t) & \text{if } m < 0 \end{cases}$$

The integer $n$ is a positive integer called the radial degree, $m$ is a negative or positive integer called the azimuthal frequency. Note that the difference between $n$ and $m$ is an even number. Hence, there are no $Z_n^m$ or $Z_n^{-m}$ Zernike polynomials. Note also that when $m$ is equal to zero, there is no angular dependency and the mode is rotationally symmetric.

The pyramidal shape of the Zernike decomposition is depicted on figure 5.

In this basis, any wavefront projection has the following expression, called horizontal decomposition:

$$P(f) = \sum_{n=0}^N \left( \sum_{m \in I_n} < f, Z_n^m > Z_n^m \right)$$

where $I_n = (-n+2\mathbb{N}) \cap (n-2\mathbb{N})$.

The orthogonality of the Zernike basis also allows to compute easily any wavefront moment, as for instance the RMS value:

$$RMS = \sqrt{\sum_{n=1}^N \sum_{m \in I_n} < f, Z_n^m >^2}$$

2.2. Interpretation of the optical aberrations in the Zernike basis

The current standard method of describing the ocular wavefront error is based on the use of the Zernike polynomial basis [4]. Each Zernike polynomial can be seen as a particular wavefront error within the eye’s pupil. Some modes, such as $Z_2^2$, called defocus, are familiar to the clinician, whereas others are not (for instance
The orthonormal basis of Zernike polynomials $Z^{-4}_4$). Table 1 gives the most common interpretation of Zernike polynomials as optical aberrations. Note that the constant polynomial $Z^0_0 = 1$ is called piston.

An the eye that has coma along axis produces a comet-shaped blur on the retina (Coma aberration comes from "comet" in latin). A blurred image occurs when light from the margin of a lens or mirror with a spherical surface comes to a shorter focus than light from the central portion, hence the name ?spherical aberration?.

\begin{table}[h]
\begin{tabular}{|c|c|}
\hline
Zernike polynomial & optical aberration \\
\hline
$Z^0_0$ & piston \\
$Z^1_1$ & vertical tilt \\
$Z^1_1$ & horizontal tilt \\
$Z^2_1$ & bent astigmatism \\
$Z^2_2$ & myopia / hyperopia \\
$Z^3_2$ & horizontal/vertical astigmatism \\
$Z^{-3}_3$ & vertical trefoil \\
$Z^{-1}_3$ & vertical coma \\
$Z^1_3$ & horizontal coma \\
$Z^{-3}_3$ & horizontal trefoil \\
\hline
\end{tabular}
\caption{The first Zernike polynomials and their corresponding eye aberration.}
\end{table}

The optical aberrations corresponding to a polynomial wavefront of maximal degree lesser or equal to 2 are the only ones that can be corrected by spectacle glasses. However, with the previous Zernike decomposition, a quasi parabolic wavefront profile, corresponding to the wavefront error caused by a myopia or hyperopia and various subtle ocular imperfections, can be wrongly interpreted, as shown on figure 6.
On this particular case, the global shape is parabolic and is due to a myopia. Firstly, the values of the order 1 coefficients in the Zernike decomposition are surprisingly not negligible (in particular, the value of the vertical tilt coefficient $z_{-1}^1$ is large). This would indicate that there exists a predominantly vertical deviation angle from the main propagation direction of the wavefront, whereas the measurements are routinely performed in a rectilinear and coaxial way. Secondly, it happens that the order 2 coefficients will not predict with a good accuracy the objective refraction, which corresponds to the quadratic wavefront deformation that can be corrected with spectacle glasses. Actually, all these artefacts in the Zernike coefficients are due to the presence of order 1 and order 2 terms in higher order Zernike polynomials, such as: $R_{13}^3(r) = 3r^3 - 2r$, $R_{14}^0(r) = 6r^4 - 6r^2 + 1$ or $R_{14}^2(r) = 4r^4 - 3r^2$. The elevation of the order 1 coefficients may be caused by the necessity to balance the order 1 terms introduced by modes such as the vertical coma ($Z_{-1}^3$). Part of the value of the order 2 coefficients are related by the presence of 2 terms which are embedded in the analytical expression of all the higher-order mode of even radial order $n$ and non-null azimuthal frequency ($m \neq 0$). It has been shown that the majority of the visual impact of high levels of fourth-order Zernike aberrations can be attributed to the second-order terms within these polynomials ([1]).

To remedy this major drawback or the Zernike basis, a new basis is presented in the next paragraph.

2.3. A new decomposition basis

For any positive integer $L$ (for instance $L = 2$), denote $(L)\mathcal{P}_{D}$ the subset of wavefronts with valuation greater or equal to $L + 1$:

$$(L)\mathcal{P}_{D} = \text{Vect}\{ X^m_n ; n \geq L + 1, m \in I_n \}$$

where, for all $n \in \mathbb{N}$, $I_n = (-n + 2\mathbb{N}) \cap (n - 2\mathbb{N})$ and for all $m \in I_n$:

$$X^m_n : (x,y) \rightarrow x \frac{n+m}{2} y \frac{n-m}{2}$$

For the chosen value of $L$, the following direct sum can be written:

$$\mathcal{P}_{D} = \mathcal{P}_{D}^{(L)} \oplus (L)\mathcal{P}_{D}$$
This direct sum, called LD-HV separation, allows to write any polynomial wavefront \( f \) into the sum of a low degree (LD) polynomial \( f_b \) (\( \text{deg}(f_b) \leq L \)) and a high valuation (HV) polynomial \( f_h \) (\( \text{val}(f_h) \geq L + 1 \)):

\[
f = f_b + f_h
\]

\( f_b \) and \( f_h \) are respectively called the low degree (respectively high valuation) component of \( f \).

The low degree component is written in the previous Zernike basis of \( P^L_D \):

\[
f_b = \sum_{n=0}^{L} \left( \sum_{m \in I_n} g^m_n(f) Z^m_n \right)
\]

For the high valuation contribution, an orthonormal basis \( \{G^m_n\}_{n \geq L, m \in I_n} \) of the subspace \((L)P^L_D \) is built with the same Gram Schmidt orthonormalization principle as previously for the definition of the Zernike basis. In this new basis, the high valuation component writes as:

\[
f_h = \sum_{n=L+1}^{\infty} \left( \sum_{m \in I_n} g^m_n(f) G^m_n \right)
\]

It can be easily seen that all the Zernike polynomials \( Z^m_n \) with azimuthal frequency higher or equal to \( L + 1 \) belong to the subspace \((L)P^L_D \). As they are already an orthonormal family, they are kept in the new basis and the orthonormalization principle is only applied to the remaining rows with index \( m \) lesser or equal to \( L \) (in absolute value). The new orthonormal basis \((L)P^L_D \) is then made of the following polynomials:

\[
\forall m \in \mathbb{Z}, \ \forall n \in (|m| + 2\mathbb{N}) \cap (L + 1 + \mathbb{N}), \quad G^m_n = \begin{cases} R^{L+1+1L+2n}_{L+1}(n) \otimes T_{(m)} & \text{if } |m| \leq L \\ Z^m_n & \text{else.} \end{cases}
\]

where the polynomials \( R^k_n \) have been previously defined and where \( 1_{L+2n}(n) \) is equal to one if \( n - L \) is an even number, and zero elsewhere.

A pyramidal description of this new LD-HV basis is given on figure 7 for \( L = 2 \).

More precisely, we get the following expression for the new polynomials of this basis:

- \( G^3_3(r \cos(t), r \sin(t)) = 2\sqrt{2} r^3 \sin(t) \)
- \( G^3_4(r \cos(t), r \sin(t)) = 2\sqrt{2} r^3 \cos(t) \)
- \( G^4_2(r \cos(t), r \sin(t)) = \sqrt{10} r^4 \sin(2t) \)
- \( G^3_1(r \cos(t), r \sin(t)) = \sqrt{5} r^4 \)
- \( G^3_2(r \cos(t), r \sin(t)) = \sqrt{10} r^4 \cos(2t) \)
- \( G^3_5(r \cos(t), r \sin(t)) = \sqrt{2} (5r^5 - 4r^3) \sin(t) \)
- \( G^2_2(r \cos(t), r \sin(t)) = 2\sqrt{2} (5r^5 - 4r^3) \cos(t) \)
- \( G^3_6(r \cos(t), r \sin(t)) = \sqrt{14} (6r^6 - 5r^4) \sin(2t) \)
- \( G^3_7(r \cos(t), r \sin(t)) = \sqrt{7} (6r^6 - 5r^4) \)
- \( G^3_8(r \cos(t), r \sin(t)) = \sqrt{14} (6r^6 - 5r^4) \cos(2t) \)

Note that in this basis, the order 0, 1 and 2 effects are now totally dissociated with all the higher order effects.

In practice, the computation of \( f_b \) et \( f_h \) for a given value of \( N \) is done first by computing the projection \( P(f) \) of \( f \) in the Zernike pyramidal basis and then the low degree and the high valuation components are obtained with a previous offline computation of each Zernike polynomial in the new basis.

Various clinical test cases are presented in the next section in order to compare the decomposition of particular wavefronts in the Zernike basis and in the LD-HV basis.
3. Clinical examples

3.1. high myopia

In this first example, low degree modes explain the major part of the wavefront error (see figure 8). As a consequence, the decomposition in the Zernike and the $LD - HV_2$ basis are grossly similar.
3.2. Myopia after laser surgery

In this second example where myopia has been treated by laser refractive surgery, the wavefront error is now mainly constituted by high-order aberrations, among which fourth order spherical aberration $Z_4^0$ dominates. After surgery, the myopic defocus has been corrected in the central area of the wavefront, but partly remains in the peripheral area because of the reduced functional zone diameter achieved by the surgery correction.
The positive defocus coefficient value residual myopia which is found in the Zernike decomposition (see figure 9) wrongly suggests the presence of a myopic refractive error. This defocus compensates for the 2nd order terms present in the analytical expression of the fourth order spherical aberration of the Zernike basis.

The difference between the second order Zernike decomposition and the corresponding contribution in the new basis is very important. On the other hand, the high value of the \( g_{44} \) coefficient confirms the predominance of positive spherical aberrations to explain most of the wavefront error persisting after refractive surgery. Note also the reduction to an almost negligible value of the other low order modes in the new basis, such as first order modes (tilt).

In order to better compare the decomposition between both bases and to quantify more precisely the low degree contribution, a new graphical representation is presented on figure 10. The purpose of this figure is to display a graphical vectorial representations of the Zernike vs \( LD - HV_2 \) classifications after surgery. The difference in the length of the low degree vector \( (G_{ld}) \) in the \( LD - HV_2 \) classification and the low degree vector \( (Z_{ld}) \) in the Zernike classification is striking (left). The sum of the low and high degree of each classification enables the reconstruction of the total wavefront \( \tilde{f} = Z_{ld} + Z_{hd} = G_{ld} + G_{hd} \) (right).

This representation strongly suggests that the new basis is more representative of the true optical disturbance that remains after surgery, which is mainly caused by high-order aberration modes, whereas the Zernike basis could lead to a misunderstanding and even a wrong diagnosis of the refractive status of the considered eye.

### 3.3. Keratoconus

The results for this new case, an eye presenting a keratoconus are depicted on figures 11 and 12. Keratoconus is characterized by a severe corneal deformation, which causes visual disturbances induced by an increase in high-order aberrations that can not be corrected by spectacles.

![Figure 11. Keratoconus. Wavefront decomposition in the Zernike and the \( LD - HV_2 \) basis.](image-url)

Note in particular on figure 11 that the prism (first order) coefficients in the Zernike decomposition are only here to compensate for the first order terms contained in the analytical expression of Zernike coma aberration modes \( (Z_{31}^{2}) \) coefficients). In the \( LD - HV_2 \) basis, the astigmatism coefficients \( G_{42}^{2} \) were confirmed to be best estimators of the refractive astigmatism measured with subjective refraction. A secondary astigmatism mode \( (G_{42}^{2} \) coefficients) is also clearly visible in the new decomposition of the wavefront error. Figure 12 also confirms
Figure 12. Keratoconus. The low degree contribution in Zernike and $LD - HV_2$ basis.

the differences between both decompositions and the relative contribution of low degree aberrations to the total wavefront. In particular, the low degree contribution to the refractive error in the Zernike decomposition is underestimated because of the presence of low degree terms which have been absorbed in its higher degree component.

4. Conclusion

A new decomposition basis of the ocular wavefront has been proposed in order to better classify and treat the optical aberrations of the human eye between their lower vs higher degree contribution. As opposed to the Zernike basis, currently used for clinical diagnosis, the new basis introduced here, called $LD - HV_L$, completely gathers the low degree terms (lesser than $L$) by purifying the remaining higher-order component of the wavefront error, at the only expense of relaxing the orthogonality condition between these two components. On various clinical examples, this new decomposition basis demonstrates its superiority by removing some undesirable artefacts of the Zernike basis, like non existing myopic defocus after myopic surgery, prismatic terms, etc. It also allows to highlight the effect of visually significant aberrations frequently encountered in the wavefront error sensing of various frequent clinical situations, such as spherical aberration, coma, and secondary astigmatism. The lack of orthogonality between the lower and higher order aberrations component may not be detrimental in clinical situations where it is important to address the consequences of the higher-order aberration modes.

References

[1] X. Cheng, A. Bradley, S. Ravikumar, and L. Thibos, Visual impact of zernike and seidel forms of monochromatic aberrations., Optom Vis Sci, 87 (2010), pp. 300–12.
[2] J. Liang, B. Grimme, S. Goelz, and J. Bille, Objective measurement of the wave aberration of the human eye using a shack-hartmann wavefront sensor., J Opt Soc Am, 11 (1994), pp. 1949–1957.
[3] B. Platt and R. Shack, History and principles of shack-hartmann wavefront sensing., J. Refractive Surgery, 17 (2001), pp. 573–577.
[4] L. Thibos, R. Applegate, J. Schwiegerling, and R. Webb, Vsia standards taskforce members. vision science and its applications. standards for reporting the optical aberrations of eyes., J. Refractive Surgery, 18 (2002), pp. 652–660.
[5] ———, The 2012 charles prentice medal lecture: wavefront measurement of refractive state., Optom Vis Sci, 90 (2013), pp. 911–23.
[6] H. Wyatt, The form of the human pupil, Vision Res, 35 (1995), pp. 2021–2036.
[7] F. Zernike, *Diffraction theory of the knife-edge test and its improved form, the phase-contrast method*, Monthly Not. of R. A. S. 94, (1934), pp. 377–384.