Improvement of dynamic features of an adaptive optical system for detection and correction of ophthalmic distortions

D. W. de Lima Monteiro, A. S. O. Furtado, O. Soloviev*, A. I. Ferreira Jr., F. B. Teixeira, J. G. M. Melo, G. Vdovin*

OptMA^{lab}, Department of Electrical Engineering, Universidade Federal de Minas Gerais, Brazil *Electronic Instrumentation Lab., Delft University of Technology, The Netherlands Av. Antônio Carlos, 6627, 31270-010 Belo Horizonte-MG, Brazil, davies@ufmg.br

ABSTRACT

Precise real-time observation of retinal features, ophthalmic biometric identification and prescription of aberration-specific contact lenses demand detection and eventual correction of ophthalmic aberrations at rates higher than those of micro-fluctuations and involuntary movements inherent to the eye. Detection and correction can be accomplished in a feedback-loop adaptive optical system comprising a wavefront sensor and a deformable mirror. The loop must operate at a frequency higher than 300 Hz, the maximum laser power entering the eye should observe safety regulations and the accuracy of aberration detection should be better than $\lambda/4$.

This article presents recent advances in application-specific sensor design, in neural-network control algorithms and in randomized wavefront sampling planes. The results also indicate the potential of standard silicon-technology components to the medical scenario.

Keywords: adaptive optics, image sensor, CMOS, wavefront sensor, ophthalmology, neural network

1 INTRODUCTION

Ophthalmic aberrations beyond defocus and astigmatism can be the culprit of poor retinal observation and unsharp vision. Optimal vision, both inwards and outwards, can only be accomplished if all aberrations in the eve optical path are corrected. This path comprises the cornea, aqueous humor, crystalline lens, vitreous humor and the retina. The appropriate distortion measurement and correction methods should be compatible with the dynamics of the eye. Even when the eye is locked at a fixed target, phenomena such as microfluctuation and drift are present as a consequence of the vestibulo-ocular reflex (VOR), that stabilizes an image on the retina, and the saccadic eye motion that maps a scene and maintains visual stimuli. These superimposed fluctuations can reach a 90 Hz frequency with image decentralization of up to 25 μ m.

Suitable adaptive optical systems, to account for dynamic changes in the eye, employ a Hartmann-Shack wavefront sensor and an adaptive membrane mirror, as

reported in [1]. A probing laser beam is reflected by the retina and traverses the ocular path. The outcoming wavefront, associated with the reflected beam, becomes imprinted with the eye distortions. The sensor output signals are related to the wavefront profile and represent the input to a control algorithm that drives the adaptive corrector (electrostatically actuated micromachined deformable mirror) [2]. The feedback loop must operate at a frequency higher than 300 Hz to allow for several iterations required for optimal convergence. Also, the maximum laser power entering the eye should be observed according to safety regulations, and the accuracy of aberration detection should be better than $\lambda/4$ [3].

Our approach addresses several system elements: mirror control method, sensor design sampling-plane architecture. We propose: a simple Adaline neural network for the control algorithm in contrast to the traditional least-square matricial method; an application-specific CMOS image sensor as opposed to an off-the-shelf camera; and controlled randomization of the wavefront sensing mask instead of a regular array of sub-apertures. The pursuit of a fast and accurate adaptive optical system by means of the aforementioned elements will be described in the following sections.

2 ADALINE NEURAL NETWORK

We focused on the very simple Adaline neural networks [4] for tests in both diagnosis of wavefront aberrations and deformable-mirror (DM) control in an Adaptive Optical (AO) system.

3 Wavefront reconstruction

In the Hartmann wavefront sensing method the probing laser beam impinges on an opaque mask with a number of sub-apertures. The beam is sampled as M subbeams, which propagate a small distance (usually from 1 to 10 cm) towards an observation screen. Assuming Fresnel diffraction does not impart too much power divergence from mask to screen and that the wavefront is flat, then, the resulting spots will unmistakenly reproduce on the screen the sub-aperture grid present on the mask. Distortions on the wavefront result in the departure of the light spots from the initial grid. Recording

the x and y displacements of the spots from the (i,j) grid points, one can geometrically calculate the respective (i,j) wavefront slopes and therefore reconstruct the initial wavefront.

A test wavefront Ξ_W can be represented on a native basis, for instance Zernike polynomials Z_i [5], as $\Xi_W = \sum_{i=0}^{\infty} C_i Z_i(\rho, \theta)$. The sensor signals, associated with the test wavefront, yield a reconstructed wavefront Ξ_R that can be decomposed on a reconstruction basis with N functions and their respective λ_i coefficients as $\Xi_R = \sum_{i=0}^{N-1} \lambda_i Z_i(\rho, \theta)$. The reconstruction order N is dictated mostly by the sensor geometry and noise.

Wavefront aberrations in the human eye have been observed to be substantial only to the first 14th Zernike orders, where the first two polynomials represent strictly non-aberrational terms such as tip and tilt, which can be related to the eye axis. This indicates that a weighted sum truncated to 14 terms might suffice for the reconstruction.

3.1 Wavefront diagnosis

For the diagnosis of test wavefronts the algorithm receives as inputs the reconstruction basis, consisting of N=14 Zernike functions. It outputs a weighted sum Ξ_R that is compared at each grid coordinate (ρ and θ) to the test wavefront Ξ_W . Then, the mean-square error (mse) of the resulting residual matrix $\Xi_W - \Xi_R$ serves as feedback to the neural net. Each internal iteration (epoch) adjusts the 14 nodal weights (λ_i) towards mse minimization. The stop parameters are mse $\leq 1e - 6$ and $epochs_{max} = 1000$. Figure 1 illustrates the network topology. The ultimate weights indicate the best coefficients (λ_i) to the respective Zernike terms in order to reconstruct the test wavefront. Data permutation in this algorithm prevents that it stacks at pseudo-minimum error levels.

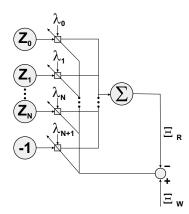


Figure 1: Adaline neural network topology.

We executed the algorithm for six different test wavefronts, represented as a weighted sum to the k^{th} order. For $k \leq N$, the test wavefront can be completely described by the set of N=14 reconstruction functions. For k>N, higher spatial frequency terms must be somehow described by a restricted range of Zernike terms, resulting in an increase in aliasing and in a worse minimum error as more high-order terms are added.

The algorithm can be easily extended to include dynamic adaptability, where the number of Zernike reconstruction functions N can be changed between epochs depending on the mse convergence rate.

3.2 Adaptive Optical system

The Adaptive Optical (AO) system was set up as shown in Figure 2. The laser beam (12-mm diameter, He-Ne, $\lambda = 632.8nm$) traverses the aberration plane (AB), impinges on the mirror and is reflected towards a beam-splitter that divides the beam between the custom CMOS wavefront sensor and a conventional CCD-based wavefront sensor. This preliminary system is intended to test the performance of components and control algorithm. It is not yet designed to incorporate an eye, its target and an observational site. However, the plane AB in the presented system can accommodate aberrations that, with no lack of generality, could be present in an eye adapted to an optical subsystem [6].

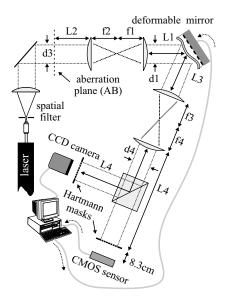


Figure 2: Adaptive Optical system.

The mirror is a 15-mm micromachined membrane deformable mirror (MMDM), Flexible Optical B.V., with 37 channels and electrostatic actuation. The membrane is kept at a bias position by maintaining an offset voltage of $\sim 160 \mathrm{V}$ at all actuators with respect to the grounded membrane. The channel voltage can reach up to 250 V and the maximum mirror stroke is 8 μ m.

The custom wavefront sensor consists of an opaque mask (Hartmann mask) with sixteen 450 μ m circular sub-apertures laid on a regular square array with

1000 μ m pitch. The mask is positioned at 8.3cm away from a CMOS chip with a corresponding array of quad-cells (QC's) 600 μ m wide [7]. Each pixel in a QC is passive and based on a double-junction structure (p+/n-well/p-epi). The custom chip outperforms an off-the-shelf camera in operational speed because it delivers signals directly proportional do the x and y displacements, circumventing the need for image processing.

The quad-cell response is non-linear and can be closely fitted with a sigmoidal curve. Although it is not detrimental to wavefront accuracy, upon calibration, linearization of this curve could speed up the numerical routines that relate the sensed signal to true displacements.

3.3 Deformable-mirror control

A mirror mode, or influence function, is the shape the membrane assumes when an incremental voltage step is applied to a single mirror electrode (actuator). For the mirror we used, there are thus 37 modes, each imparting a characteristic set of spot displacements on the CMOS wavefront sensor. We can consider the wavefront reflected from a biased membrane to be the reference wavefront, itself responsible for a set of x and y displacements marking reference positions.

The Adaline neural network used for Zernike terms has been modified to attend to this system. The topology remains basically the same, but the input reconstruction functions have been substituted with the sensor responses to the 37 mirror modes M_i . Each mode corresponds to a vector with 32 elements representing 16 x and y displacements. The algorithm delivers an output vector D, as a weighted sum of the basis vectors, in an attempt to minimize the error to a sample vector D_d associated with the aberrated wavefront Ξ_W . The weights w_i (i=0,...,36) are related to the voltages to be applied to the mirror actuators.

3.4 Algorithm performance

We introduced centered defocus with an amplitude of 36 μ m (> 55 λ) as the aberration. This value is far beyond the working stroke of the mirror membrane and is only intended to evaluate the convergence ability of the algorithm. This aberration exploits the spatial dynamic range of the quad-cells because it forces the outer spots to the edge of their respective QCs. The network parameter mse indicates how close the solution is to the reference grid as a function of the number of epochs, and the displacement error associated with it is kept under 1% for all epochs.

Next we introduced a milder aberration with amplitude $\lambda/2$ (Figure 3) to check the algorithm convergence.

When the neural algorithm starts with unity weights it takes a single epoch to reach an acceptable solution,

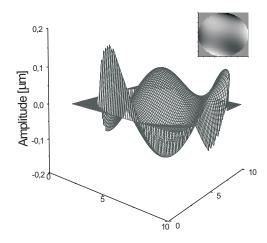


Figure 3: Profile of the mild aberration introduced, which contains mostly coma, astigmatism, trefoil and spherical aberration.

i.e. a wavefront match better than $\lambda/50$. After ~ 80 s and 1.5M epochs the mse reaches 0.00017 (wavefront match better than $\lambda/5000$).

In practice, dynamic distortions of a given wavefront occur smoothly and the network weights oscillate within a somewhat limited range. By modifying the wavefront only slightly ($\sim \lambda/16$ rms) we registered the algorithm convergence in a single epoch to mse=0.0018 (accuracy $\sim \lambda/1000$), in 200 μ s.

As the sensor has been read at 1kHz (1ms) and all software routines take less than 400 μ s the system has potential to be operated at rates higher than 600 Hz.

A single iteration of the AO system restores the wavefront to approximately $\lambda/5$ rms deviation from the reference wavefront, which complies to the wavefront accuracy of $\lambda/4$ usually demanded for ordinary eye tests, where dominant low-order aberrations might be as high as 5λ [9]. The rate of accuracy improvement with increasing iterations is yet to be investigated.

To improve the accuracy we need to employ a more accurate reference wavefront [10] and a larger number of quad-cells, which might lower the system operational frequency. However, even if it operates at 300 Hz it still offers time for 3 iterations whereas maintaining compatibility with the micro-motion frequency of the eye.

4 NEW PIXEL STRUCTURE

The CMOS sensor previously described renders a poor wavefront accuracy ($< \lambda/3$) for the low light levels demanded by ophthalmological applications. According to the American National Standards Institute (ANSI, Z-136.1 standard) [9], the maximum permissible exposure

(MPE) for 25 μ s, at $\lambda = 0.63 \mu$ m, is 1.78mW. Considering that only about 1% of the incident light is reflected by the retina, a very faint beam will reach the wavefront sensor, where each resulting light spot takes up less than about 0.5μ W.

To comply to these margins, the photosensitive structures in each quad-cell must have an improved signal-to-noise ratio (SNR), which can be achieved through noise reduction and enhanced quantum efficiency (QE). These figures directly impact the positional signal-to-noise ratio (SNR $_p$) related to the quad-cell response, which consequently affects the wavefront accuracy.

Therefore, we devised a photodiode in standard 1.6- μ m nwell CMOS technology, where some design issues were attempted in order to improve the above figures. The photodiode perimeter is that of a quarter-circle, 200 μ m radius. The main junction is n-well/p-epilayer, 2.9 μ m deep to promote maximum absorption at $\lambda=0.63~\mu$ m. The n-well is segmented to improve the junction area along with the volume of the depletion region. The region closest to the right-angle corner features and extra-shallower junction (p+/n-well) to improve sensitivity at the central part and linearize the response of the envisioned quad-cell.

The measured NEP (noise-equivalent power, SNR = 1) is $0.04\mu\mathrm{W}$. In the optical power range from $0.1\mu\mathrm{W}$ and $0.5\mu\mathrm{W}$ the SNR ranges from 5.1dB to 20.5dB, the photo-generated current I_p and the respective rms noise range from 30nA (17.8nA) to 190nA (18.3nA). The overall quantum efficiency (QE) was measured to be 71.3%. The total photodiode capacitance is 4.47pF (10 times lower than that of the previously used photodiode).

The photodiode was further equipped with 4 NMOS FETs in an active-pixel configuration, as shown in Figure 4). This architecture decouples the pixel capacitance from the signal-line capacitance and guarantees reduced reset noise and fixed-pattern noise. The reduced equivalent capacitance also reduces the signal delay time [11].

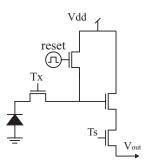


Figure 4: Active pixel with 4 FET transistors.

The signal-to-noise ratio of the active pixel is far better than that of the photodiode alone, ranging from 35dB to almost 50dB in the range from from $0.1\mu\mathrm{W}$ and $0.5\mu\mathrm{W}$. The measured active-QC response to a $0.2\mu\mathrm{W}$ spot displacement in the x direction is shown in Figure 5. It shows the linearization of the passive-QC curve. The active QC achieved a position resolution as low as $2\mu\mathrm{m}$ ($0.2\mu\mathrm{W}$ spot), indicating a wavefront accuracy of approximately $\lambda/50$. These results are undoubtedly compatible with the low-light-level ophthalmic standards.

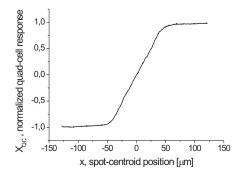


Figure 5: Response of a quad-cell with four active pixels to a 0.2μ W spot displacement in one direction.

5 RANDOMIZATION OF THE HARTMANN MASK

The performance of Hartmann masks with random position of subapertures (s/a) was recently investigated in [12]. It was shown that for Kolmogorov atmospheric turbulence randomization of s/a centers positions improves the performance of HS wavefront sensor in terms of speed and accuracy. Despite the absence of comprehensive statistics on human-eye aberrations, some advantages of randomized arrays can be demonstrated for use in ophthalmology.

5.1 Mathematical model of HS wavefront reconstruction

Mathematically, the modal wavefront reconstruction with HS test can be represented as a result of some linear operator \mathcal{H} applied to the impinging wavefront f, i.e. $\tilde{f} = \mathcal{H}f$, where \tilde{f} is the reconstructed wavefront. The operator \mathcal{H} is closely related to the geometry of the Hartmann mask and the algorithm used for wavefront sensing and reconstruction (we do not consider here gradient-measurement errors , e.g. due to non-uniform illumination or camera spatial discretization). Suppose decomposition modes $f_j, j=1\ldots, N$ and "native" basis of the incoming wavefront $g_i, i=1,\ldots,\infty$ are used, that is

$$\tilde{f} = \sum_{j=1}^{N} \lambda^{j} f_{j}, \quad f = \sum_{j=1}^{\infty} c^{j} g_{i} + c^{0}.$$

Then the operator \mathcal{H} is fully characterized by its matrix \mathbf{H} , an $N \times \infty$ matrix which j-th column is given by the coefficients of reconstruction of a wavefront containing only j-th basis function:

$$\mathcal{H}f = (f_1, f_2, \dots, f_N) \cdot \mathbf{H} \cdot \begin{pmatrix} c^1 \\ c^2 \\ \dots \end{pmatrix}. \tag{1}$$

For an often case of the same decomposition and native basis $f_j = g_j$, the ideal matrix \mathbf{H} is a unity matrix for its first N columns and zero elsewhere; a Hartmann sensor with such a matrix reconstructs the first N aberrations without distortion and is not sensitive to higher-order terms:

$$\mathcal{H}_{\text{ideal}}\left(\sum_{i=1}^{\infty} c^{i} f_{i} + c^{0}\right)$$

$$= (f_{1}, \dots, f_{N}) \cdot \begin{pmatrix} 1 & 0 \\ & \ddots \\ & & 1 \end{pmatrix} 0 \dots \end{pmatrix} \cdot \begin{pmatrix} c^{1} \\ c^{2} \\ \dots \end{pmatrix}$$

$$= c^{1} f_{1} + c^{2} f_{2} \dots + c^{N} f_{N}. \quad (2)$$

For non-ideal matrices, the aliasing and cross-coupling errors decrease the fidelity of the reconstructed wavefront.

Equation 2 allows one to estimate the reconstruction error $(\varepsilon = \tilde{f} - f)$ for any given impinging wavefront f; more practical is to derive statistical characteristics of ε from those of the coefficients c_i of sensed aberrations, defined in terms of average vector $\langle c_i \rangle$ and correlation matrix $\langle c_i c^j \rangle$.

In the absence of statistical data, no definitive statement about the advantages of a particular mask geometry can be done. However, some quality criteria such as the similarity to an ideal matrix may help evaluate the performance improvement. In the next section, rms value of deviation of \mathbf{H} from ideal matrix \mathbf{H}_{ideal} is used as a quality factor Q:

$$Q = \text{RMS}(\mathbf{H} - \mathbf{H}_{\text{ideal}}) = \sqrt{\frac{\sum_{i,j} (h_i^j)^2 - N}{NM}}, \quad (3)$$

where h_i^j is an element of **H** and number M of columns in **H** used for calculation is large enough. Q represents an average coefficient in **H**, magnitude of "feed-through" of high-order modes to each of the first N modes; the larger Q, the greater aliasing error is; for the ideal Hartmann operator, Q = 0.

The following example gives an illustration of the Q-factor. Consider the first N modes as a signal contaminated by hypothetical "white Zernike noise", a random aberration with zero time average and a cross-correlation matrix $\langle c_i c^j \rangle = \sigma_{\text{noise}}^2 \delta_i^j$, where δ_i^j is the Kronecker symbol. If the bandwidth of the system is

M modes, the noise total power is $M\sigma_{\text{noise}}^2$. According to Equation (11) from [12] and Equation (3) from the previous paragraph, this results in aliasing error with total power $sigma_{\text{al}}^2 = Q^2NM\sigma_{\text{noise}}^2$. Thus SNR of reconstructed and incoming wavefront are related as

$$SNR_{r} = \frac{1}{\sqrt{NQ}}SNR_{i}.$$
 (4)

5.2 Numerical results

We have considered 3 mask geometries shown in Figure 6 with a total aperture diameter of 12 mm and s/a diameter $450\mu m$: one regular 64-hole array and two random arrays with 64 and 32 sub-apertures. The pitch of the regular grid is 1 mm; the same value was used as minimal distance between s/a centers in the randomized arrays to maintain the dynamical diapason.

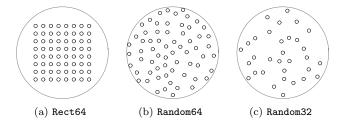


Figure 6: Regular 64-hole Hartmann mask, randomized 64-hole and 32-hole masks used for calculations.

Matrices **H** calculated for these three geometries for Zernike polynomials as both decomposition and native basis, N=14 decomposition modes, and least-squares decomposition algorithm (see [12] for details on the computational procedure). The matrix for the regular array contains larger elements than that for the randomized mask with 64 s/a, but has sparser structure. The matrix for the 64-hole randomized mask appears to be the closest to the ideal matrix in the least-square sense. For hypothetical white Zernike noise, sensing with mask Random64 will decrease the signal-to-noise ratio in ≈ 2.1 times with respect to ≈ 2.9 for the regular mask; or one can use a random mask with half the number of s/a, which keeps the aliasing error at the same level, whereas it speeds up the calculations in 4 times.

In practice, the spectrum of aberrations usually decreases with the aberration order. Thus, the main aliasing error for a wavefront reconstructed with the first 14 modes is expected to come from the next several aberrations. Here the randomized masks are superior to the regular arrays. Figures 7(a)–(c) show a close-up view of the matrices \mathbf{H} ; please note the uneven aliasing for the regular array leading to larger rms error.

Random arrays allow for another performance improvement by using a special trick, reported here for the first time, to our knowledge. Namely, being interested in

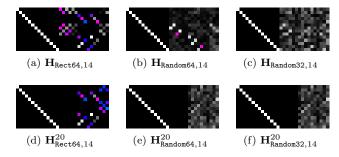


Figure 7: (a)–(c): first 30 columns of matrices of Hartmann operators **H**; (d)–(f): first 30 columns of first 14 rows of matrices of Hartmann operators **H** calculated for 20 decomposition modes.

the first N terms, one is able to remove *completely* aliasing from the next n terms by calculating a reconstructed wavefront for N+n terms and then discarding unnecessary coefficients. This is equivalent to considering only first N rows of **H** built for N + n decomposition modes. This technique works only for random Hartmann masks, mainly due to the dramatic growth of aliasing error for regular arrays [12]. For regular masks, matrices **H** are uneven, with larger elements concentrating in the first rows; discarding last rows, with smaller values, one increases the average element (Q-factor). For randomized Hartmann masks the matrices are nicely uniform, so the average remains at the same level after discarding the last rows. This is illustrated by the examples shown in Figure 7(d)–(f) and by the Q-values in Table ??; the matrices $\mathbf{H}_{\mathtt{mask},14}^{20}$ were obtained by discarding the last 6 rows from the matrices calculated for 20 decomposition rows.

6 CONCLUSIONS

We presented the initial tests indicating the feasibility of using a simple and yet robust neural network for both the diagnosis of human-eve aberrations and the control of a micromachined deformable mirror in an Adaptive Optical system. The algorithm performs adequately in terms of both speed and accuracy. It also offers the possibility for real-time adaptability. To comply to the low-light levels demanded in ophthalmology, we proposed, designed and tested an alternative pixel structure for use in an active quad-cell. The quad-cell features a position resolution that yields a wavefront accuracy as good as $\lambda/50$ for a $0.2\mu W$ spot; it operates satisfactorily even for spots five times fainter. At last, we studied the randomization of sub-apertures positions on a Hartmann mask and concluded that a random mask leads to a substantial anti-aliasing effect if proper design and algorithm boundaries are observed.

ACKNOWLEDGMENTS

The projects presented herein have been in various ways supported by FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais), PRPq/UFMG, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and by the Dutch Technical Foundation (STW). The authors thank the following people for their unconditional practical assistance Mr. H. Lopes Santos, Mr. L. M. de Aquino Macedo, Mr. H. A. Rodrigues (UFMG), Ing. W. van der Vlist and Ing. J. Bastemeijer (DIMES), and Dr. M. Loktev (OKO Technologies).

REFERENCES

- E.J. Fernandez, P. Artal., "Study on the effects of monochromatic aberrations in the accommodation response by using adaptive optics," J. Opt. Soc. Am., 1732-1738, 2005
- [2] G. Vdovin et al., "Technology and applications of micromachined silicon adaptive mirrors," Opt. Eng. 36, p.1382-1390, 1997
- [3] D. Droste and J. Bille, "An ASIC for Hartmann-Shack wavefront detection," IEEE Journal of Solid State Circuits, **37**, 173-182 (2002)
- [4] B. Widrow and M.E. Hoff, Jr. "Adaptive switching circuits." Western Electric Show and Convention Record Part 4, p.96-104, 1960
- [5] M. Born, E. Wolf., "Principles of Optics," Pergamon Press, Oxford, 6th edition, 1986.
- [6] G. Vdovin et al., "Adaptive correction of humaneye aberrations in subjective feedback loop," Opt. Letters 30/7, 2005
- [7] D. W. de Lima Monteiro et al., "High-speed wavefront sensor compatible with standard CMOS technology," Sensors and Acutators A, 109 (3),p.22-230, 2004
- [8] G. Vdovin., "Adaptive mirror micromachined in silicon," Delft University Press, Delft, 1996
- [9] J. Liang., "A new method to precisely measure the wave aberrations of the human eye with a Hartmann-Shack-Wavefront-Sensor," PhD Thesis, University of Heidelberg, 1991
- [10] G Vdovin., "Optimization-based operation of micromachined deformable mirrors," Proc. SPIE, 3353, 902-909, 1998
- [11] D. W. de Lima Monteiro., "Fast Hartmann-Shack wavefront sensors manufactured in standard CMOS technology," 5/5, 1530-437, 2005
- [12] O. Soloviev and G. Vdovin, "Hartmann-Shack test with random masks for modal wavefront reconstruction," Opt. Express 13, 9570-9584 (2005) http://www.opticsinfobase.org/ abstract.cfm?URI=oe-13-23-9570