ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

Neural Network Performances in Astronomical Image Processing

Rossella Cancelliere, Mario Gai *
Dipartimento di Matematica, Università di Torino

v. C. Alberto 10, 10123 Torino, Italy

Abstract. In this paper we use neural networks to verify the similarity of real astronomical images to predefined reference profiles. We use an innovative technique to encode images that associates each of them with its most convenient moments, evaluated along the $\{x, y\}$ axes; in this way we obtain a parsimonious but effective method with respect to the usual pixel by pixel description.

1 Introduction

The multilayer perceptron, the first largely successful neural network model, has been largely used for a great variety of applications after its presentation by D. Rumelhart et al. [3]; recently some attempts to use neural networks in astronomy have been performed, mainly in the field of adaptive optics: the reader can find details in the papers by Lloyd-Hart et al. [2] and Wizinowich et al. [4].

The problem we want to face is that of verifying the compatibility of the real images with some reference profiles; also it's really important the capability of extracting from the data some parameters suitable for a new definition of the template, in order to improve its similarity to the data. Self-calibration of the data, by deduction of the parameters for optimisation of the image template, is a key element in the control of the systematic effects in the position measurement.

Our objective is the implementation of a tool for analysis of realistic images and deduction of a set of aberration parameters able to describe their discrepancy with respect to the ideal, non-aberrated image; this is obtained using a sigmoidal neural network with one hidden layer, trained by the usual backpropagation algorithm and training and test sets derived from opportunely codified instances.

*Osservatorio Astronomico di Torino, str. Osservatorio 20, 10025 Pino Torinese, Torino, Italy

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

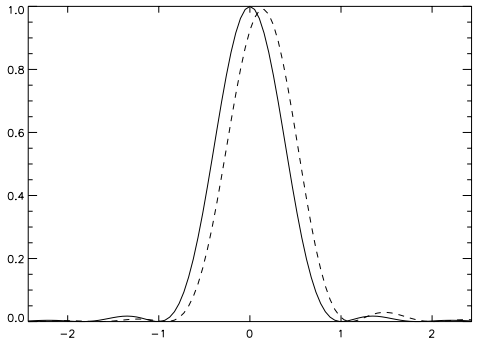

Figure 1: Airy image (solid line) and aberrated image (dashed line).

The paper is structured in three sections: in Section 2 we discuss the image characterisation problem addressed in the present work, both in a general case and in the simplified framework adopted here. In Section 3 we describe the data set generation, its processing and the current results.

2 Imaging in Astronomy

The Fraunhofer diffraction theory, described in many textbooks on optics (see [1], from which the notation is derived), is the basic framework in which to develop our problem. We only remind here that the ideal monochromatic image of a point-like source at wavelength λ, obtained from an unobstructed circular pupil of diameter D, without aberrations, is radial and described by the Airy function:

$$I\left(r\right) = k\left[2\,J_1\left(r\right)/r\right]^2\,, \tag{1}$$

where J_1 is the Bessel function of the first kind, order one, and $r = D/2$ is the aperture radius; the Airy function is shown by the solid line in Fig. 1, compared with an aberrated image (dashed line) affected by one wavelength of coma (see below). The Airy radius, enclosing the central lobe, is $1.22\lambda/D$.
In a simplified framework the perturbations introduced on the image by real optics can be described as a perturbation of the wavefront by the five classical Seidel aberrations.
The wavefront deviation from planarity is expressed by the phase aberration Φ, which defines the pupil function $e^{i\Phi}$, and the diffraction image on the focal plane by the square modulus of the Fourier Transform:

$$I\left(r,\phi\right) = \frac{k}{\pi^2}\left|\int_0^1 d\rho \int_0^{2\pi} d\theta\,\rho\, e^{i\Phi(\rho,\theta)} e^{-i\pi r\rho\,\cos(\theta-\phi)}\right|^2 \tag{2}$$

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

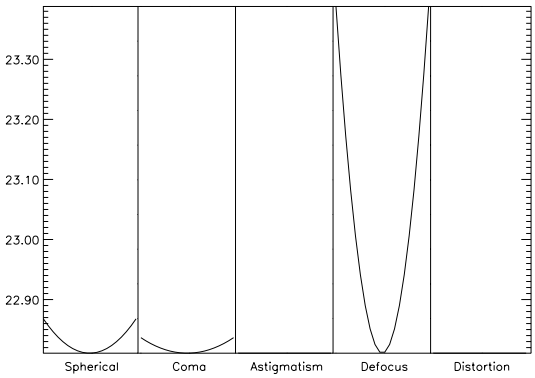

Figure 2: Variation of the x variance vs. each aberration.

where ρ and θ are the pupil coordinates (normalised radius and azimuth), whereas r and ϕ are the image coordinates. If $\Phi = 0$ (non-aberrated case), eq.(1) is retrieved.

In the simplified framework of the five classical Seidel aberrations, i.e. A_s: Spherical aberration, A_c: Coma, A_a: Astigmatism, A_d: Defocus (field curvature) and A_t: Distortion (t = tilt), the phase aberration is generated by superposition of the five terms:

$$\Phi(\rho,\theta) = \frac{2\pi}{\lambda}\left[A_s\rho^4 + A_c\rho^3\cos\theta + A_a\rho^2\cos^2\theta + A_d\rho^2 + A_t\rho\cos\theta\right]; \quad (3)$$

by replacement of eq.(3) in eq.(2), we see that the relation among the aberrations and the image is nonlinear.

We consider the regime of "small" or "acceptable" aberrations, corresponding to the classical Rayleigh criterion of one quarter of wavelength ($[-0.25, 0.25]$), for the test set. We select for training a slightly larger interval $[-0.3, 0.3]$, to avoid potential boundary problems. Over a larger aberration range, the image quality degrades significantly, and we are most interested in network classification capability in the case of small image perturbations. The Seidel aberrations are the targets of our network, therefore the output vector of the neural network is five-dimensional, whereas the encoding of the input image is discussed below.

2.1 Image encoding

The encoding scheme we adopt for the images allows extraction of the desired information for classification, without resorting to the expensive classical pixel by pixel technique. Each input image is described by the centre of gravity and the first central moments, up to the fourth order:

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

$$\mu_x = \frac{\int dx\,dy\; x \cdot I(x,y)}{\int dx\,dy\; I(x,y)} \qquad \mu_y = \frac{\int dx\,dy\; y \cdot I(x,y)}{\int dx\,dy\; I(x,y)}$$

$$\sigma_x^2 = \frac{\int dx\,dy\; (x-\mu_x)^2 \cdot I(x,y)}{\int dx\,dy\; I(x,y)} \qquad \sigma_y^2 = \frac{\int dx\,dy\; (y-\mu_y)^2 \cdot I(x,y)}{\int dx\,dy\; I(x,y)} \qquad (4)$$

$$M(i,j) = \frac{\int dx\,dy \left(\frac{x-\mu_x}{\sigma_x}\right)^i \left(\frac{y-\mu_y}{\sigma_y}\right)^j \cdot I(x,y)}{\int dx\,dy\; I(x,y)}$$

The central moments are much less sensitive than the image itself to the effects related to the finite pixel size; besides, they can be deduced also from the low resolution images, without the need for high resolution detectors. The central moments have an immediate physical meaning:

- the first order moment provides the centre of gravity of the image
- the second order central moment is the mean square width
- the third order central moment (skewness) is an index of the image asymmetry
- the fourth order central moment (kurtosis) is an index of how much peaked is the distribution; for a Gaussian, its value is 3

We have now to verify the sensitivity of these quantities to aberrations; for instance the x square width is shown in Fig. 2, where the values obtained for independent variation of each aberration, are plotted side by side for ease of comparison over the selected range $[-0.3\lambda, 0.3\lambda]$. It appears to be sensitive mostly to defocus, much less to spherical aberration and coma, and insensitive to astigmatism and distortion. Similarly, we evaluate the behaviour of the other low order moments with respect to variations of each aberration, selecting the suitable input variables for our neural network. The ten selected moments are σ_x^2, σ_y^2, $M(0,3)$, $M(0,4)$, $M(1,1)$, $M(2,1)$, $M(1,2)$, $M(3,1)$, $M(1,3)$ and $M(2,2)$, evaluated in two different image points. Therefore, the input vector to the neural network has 20 components.

3 Results

In this section we describe in more detail the generation of the training and test sets and the results of neural network evaluation. As described above, each input/output instance is represented by the 20 moments plus the five aberrations. The image is built accordingly to eq.(2) and (3), starting from each set of five aberration values, and used to compute the corresponding moments, according to eq.(4).

The training set is a list of $T = 2500$ instances. We use a denser distribution of the points along the individual axes of the target space, placing 100 samples

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

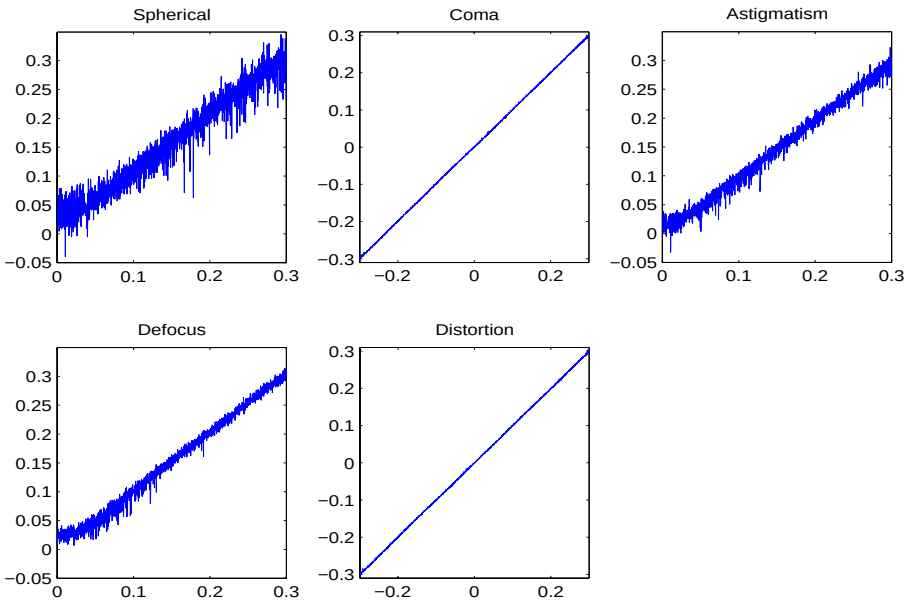

Figure 3: Neural Network Performances.

on each axis (variation of a single aberration at a time); the remaining 2000 points are randomly distributed, i.e. each variable is independently chosen.

The test set is built in a similar way, picking independent variation for each output variable, building the image and evaluating the input variables to the network. A total of 2500 test instances has been generated.

A sigmoidal neural network with an hidden layer made by 100 units was optimised on the training set and verified on the test set; the training required 8000 iterations. Because the desired behaviour of the neural network is the computation, over the test set, of output values coincident with the pre-defined target values, the plot of output vs. target, shown in Fig. 3, should be ideally a straight line ($y = a + bx$) at angle $\pi/4$, passing for the origin, ($a = 0$, $b = 1$). We compute the best fit parameters, their errors, the RMS discrepancy σ of the network output with respect to the targets, and the percentage of test instances within the confidential interval $\pm 3\sigma$; the results are shown in Tab. 1.

We remark that the outputs computed from the network are quite consistent with the desired test targets; coma and distortion are the best recognized aberrations, whereas spherical aberration, astigmatism and defocus are affected by larger systematic errors and larger dispersion.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 515-520

	$Linear fits$		σ	$\pm 3\sigma$
Spherical	$a = 0.0126$ $b = 0.975$	$\sigma_a = 0.0006$ $\sigma_b = 0.004$	0.016	97.8
Coma	$a = 0.00069$ $b = 0.9990$	$\sigma_a = 0.00002$ $\sigma_b = 0.0001$	0.0012	98.6
Astigmatism	$a = -0.0009$ $b = 0.986$	$\sigma_a = 0.0003$ $\sigma_b = 0.002$	0.008	98.4
Defocus	$a = 0.0024$ $b = 1.004$	$\sigma_a = 0.0003$ $\sigma_b = 0.002$	0.008	98.1
Distorsion	$a = -0.00030$ $b = 0.9983$	$\sigma_a = 0.00002$ $\sigma_b = 0.0001$	0.001	99.4

Table 1: Network output evaluation: linear fit of output vs. target, RMS discrepancy (σ) and percentage of results within 3σ.

4 Conclusions

In this paper we use a neural network to reconstruct aberrations in astronomical images. We test both the possibility of encoding the problem in a compact set of image descriptors with an immediate physical meaning, i.e. the moments, and the performance of the network.
We achieve good results, confirming that the problem can be solved by this technique.

References

[1] M. Born and E. Wolf. *Principles of optics*. Pergamon, New York, 1985.

[2] M. Loyd-Hart, P. Wizinowich, B. McLeod, D. Wittman, D. Colucci, R. Dekany, D. McCarthy, J.R.P. Angel, and D. Sandler. First results of an on-line adaptive optics system with atmospheric wavefront sensing by an artificial neural network. *ApJ*, 390:L41–44, 1992.

[3] D. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representation by error propagation. In *Parallel Distribuited Processing (PDP): Exploration in the Microstructure of Cognition*, volume 1, pages 318–362. MIT Press, Cambridge, Massachussetts, 1986.

[4] P. Wizinowich, M. Loyd-Hart, and R. Angel. Adaptive optics for array telescopes using neural networks techniques on transputers. In *Transputing '91*, volume 1, pages 170–183. IOS Press, Washington D.C., 1991.