EFFICIENT COMPUTATION AND NEURAL PROCESSING OF ASTROMETRIC IMAGES

Rossella Cancelliere
Department of Computer Science
University of Torino
c. Svizzera 185
10149 Torino, Italy
e-mail: cancelli@di.unito.it

Mario Gai
Istituto Nazionale di Astrofisica
Osservatorio Astronomico di Torino
V. Osservatorio 20
10025 Pino T.se (TO), Italy
e-mail: gai@oato.inaf.it

Manuscript received 21 December 2007; revised 4 July 2008
Communicated by Vladimír Kvaček

Abstract. In this paper we show that in some peculiar cases, here the generation of astronomical images used for high precision astrometric measurements, an optimised implementation of the DFT algorithm can be more efficient than FFT. The application considered requires generation of large sets of data for the training and test sets needed for neural network estimation and removal of a systematic error called chromaticity. Also, the problem requires a convenient choice of image encoding parameters; in our case, the one-dimensional lowest order moments proved to be an adequate solution. These parameters are then used as inputs to a feed forward neural network, trained by backpropagation, to remove chromaticity.

Keywords: Fourier transform, image processing, neural network diagnosis

Mathematics Subject Classification 2000: 42, 68, 85
1 INTRODUCTION

The European Space Agency has approved the mission Gaia [1], aimed at a high precision astrometric survey of our Galaxy, for launch in 2011; the Announcement of Opportunity for the data reduction has been issued on November 9, 2006. The mathematical and computing tools used for modern astronomical experiments must meet challenging requirements on resolution and precision, consistent with the measurement goals.

In particular, the location of the diffraction image of an object measured by a real instrument is affected by an apparent displacement, dependent on the source spectral distribution; this effect is called chromaticity [2]; it is described from a mathematical point of view below and clarifications about the state of the art of its effects and about the applied techniques to remove it can be found in [3]. Besides, the formulation of the diffraction integral leads naturally to its implementation based on the Fourier transform (FT).

The image of a star, considered as a point-like source at infinity, and produced by an ideal telescope, with focal length $F$, an unobstructed circular pupil of diameter $D$, at wavelength $\lambda$, has radial symmetry and is described by the Airy function (1) (see [4] for notation).

$$I(r) = k \left| 2 J_1(\pi r D / \lambda F) \right|^2.$$  

(1)

Here $J_1$ is the Bessel function of the first kind, order one, $k$ a normalisation constant, and $r$ the radial coordinate on the focal plane. The image has a characteristic size (the Airy diameter) $2.44 \lambda / D$.

The use of real telescope, however, produces diffraction images (hereafter called real images) described by the square modulus of the FT of the pupil function $e^{i \Phi}$ and characterized by a set of aberration values that depend on the phase aberration $\Phi$:

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

(2)

where $\{r, \phi\}$ and $\{\rho, \theta\}$ are the radial coordinates on image and pupil plane, respectively, and the integration domain corresponds to the pupil: for the circular case, $0 \leq \rho \leq 1$; $0 \leq \theta \leq 2\pi$. In case of a rectangular pupil, it is more convenient to use Cartesian coordinates on both image and pupil plane, e.g. $\{x, y\}$ and $\{\xi, \eta\}$ integrated between the appropriate boundaries.

The phase aberration $\Phi$ describes for the real case the deviation from the ideal flat wavefront, i.e. the wavefront error (WFE), and is usually decomposed by means of the Zernike functions $\phi_n$ ([4]):

$$\Phi(\rho, \theta) = \frac{2\pi}{\lambda} \text{WFE} = \frac{2\pi}{\lambda} \sum_{n=1}^{21} A_n \phi_n(\rho, \theta).$$  

(3)

If $\Phi = 0$ (non-aberrated case, $\{A_n\} = 0$), we obtain a flat wavefront, i.e. WFE = 0, and Equation (1) is retrieved for the circular pupil.
The WFE itself is independent from wavelength, but wavelength dependence in the pupil function is induced by the $2\pi/\lambda$ factor. Also, the nonlinear relationship between the set of aberration coefficients $A_n$ and the image is put in evidence by replacement of Equation (3) in Equation (2).

The FFT algorithm associates the resolution in one of the domains to the full range of the corresponding conjugate variable; besides, the diffraction integral is physically limited to the real pupil size. Therefore, the simplest implementation generates just two points over the characteristic length of the system, which is clearly not sufficient to provide an acceptable image detail. A typical approach to overcome this limitation consists in a formal extension of the pupil domain to a much larger size, to achieve the desired resolution, provided the function used as argument is set to zero in the extended interval. Besides, this approach involves a significant amount of computation, required from the formal definition, over quantities set to zero. This consideration led us to test the alternative approach of direct computation of the DFT in its original formulation derived from discretization of the Fourier integral.

The number of points used to evaluate the discrete Fourier transform by FFT is, however, crucial, because of its impact on both resulting precision and required computational effort; in this paper we present a method that allows usage of a number of points suitable to the desired resolution, reducing at the same time the computational cost with respect to the classic FFT algorithm.

Because the real polychromatic image of an unresolved stellar source is produced by integration over the appropriate bandwidth of the monochromatic PSF above, it is evident that objects with different spectral distributions have different image profiles; this fact implies that the position estimate produced by any location algorithm (e.g. the centre of gravity, COG), evaluated on the image profile itself, is affected by discrepancy with respect to the nominal position generated by an ideal optical system.

The location estimate is sensitive to the algorithm used [5], in case of discrepancies between the real signal and its expected profile, so that it is crucial to model the instrument and maintain its calibration with high precision. The chromaticity can be minimized by instrument design and construction [2], but the residual chromaticity must be taken into account in the data reduction phase. In this paper therefore our aim is the diagnosis of chromaticity that affects images which are realistic in the sense of a signal profile variation compatible with our expectations on the imaging quality of the instrument. Besides, at this stage we do not address the effects of random noise, since we feel necessary to identify clearly the limiting performance with respect to systematic errors. The realistic range of image variation is achieved by selection of the aberration range (in Section 3); the PSFs are not affected by readout or photon noise.

In past works [6] we developed a method for Seidel aberration estimate from the focal plane images, but the greater number of aberration terms to be detected in the Zernike decomposition would require a huge increase in the data set of examples for proper training, making the computational effort unbearable. Therefore, here we study how to identify chromaticity from the image profile itself, using
the diagnosis capability of a neural network (NN) properly trained. Our choice is
due to the existence of different works concerning the use of NN in astronomical
adaptive optics (see [7, 8]) where NN always demonstrated great robustness to noise
and damages and powerful capabilities of flexible learning.

Real-world image processing systems frequently represent a chain of hierarchi-
cally organized, interacting components ranging from basic preprocessing to high-
level image analysis and interpretation. Functional operations such as preprocess-
ing, feature extraction, data reduction/compression segmentation, object recogni-
tion, image understanding, and scene analysis have to be applied to different structural
levels of data complexity ranging from pixel data, local features, structure and
texture level data to objects, object arrangements, scene and context description.
Neural networks, as a special kind of learning and self-adapting data processing
system, have to offer considerable contributions to this field. Their abilities to han-
dle noisy and high-dimensional data, nonlinear problems, large data sets etc. have
led to a wide scope of successful applications in digital image processing. A very
interesting survey on this theme can be found in [9].

Other interesting works can be found in literature about neural network capabil-
ities to solve diagnosis, recognition or classification tasks in the framework of image
processing: problems concerning character recognition ([10]), image processing in
medical applications ([11]), image compression ([12]), and so on.

A very interesting application concerns the use of cellular neural networks
(CNN); CNN are members of the hardware family called vision chips. Based on
state-of-the-art technology, a vision chip is defined as a VLSI chip that can perform
image processing tasks. The theory of CNN develops in two main fields: cellular au-
tomata and neural networks; as an interdisciplinary product, CNN utilizes cellular
hardware structures to gain ultrahigh image processing speed ([13]).

The paper is organized as follows: in Section 2 we discuss the Fourier trans-
form computation, comparing the proposed technique with the classical fast Fourier
transform algorithm, and we present the image encoding method. In Section 3 we
describe the generation of the data sets and in Section 4 we resume the main fea-
tures of sigmoidal NN and backpropagation algorithm, with a brief reminder of the
specific definitions, then we discuss the data processing and the obtained results.

2 FOURIER TRANSFORM COMPUTATION
AND DIFFRACTION IMAGE ENCODING

In this section we describe the generation of the Fourier transform and the identifi-
cation of convenient parameters for encoding of realistic images (as expected from
real telescopes), according to Equation (2).

2.1 Discrete Fourier Transform

There are many science applications for which an accurate Fourier transform of the
signal is necessary. For example, to achieve the frequency spectrum $X(\omega)$ of a time
3 DATA SETS GENERATION

In this section we describe the generation of the training and test sets and the identification of the most convenient image parameters.

With respect to our previous work on chromaticity (see [3]), the current simulation takes into account a number of realistic instrument contributions, although some still in a simple form. In general, such terms have the effect of reducing the image sharpness, which in turn reduces both the chromaticity and the signature on the moments. It is therefore important to verify whether the principle of chromaticity estimation from the moments still holds, also in the case of more realistic images, and to evaluate the impact on the performance. In practice, the relationship between chromaticity and some critical moments might be degraded significantly, or the number of cases required for proper NN training might increase to unmanageable levels.

![Chromaticity plots](image)

**Fig. 4.** From top to bottom: distribution of chromaticity vs. image centre of gravity, RMS width, third and fifth order moments

We verify that the relationship between chromaticity and moments remains good (see Figure 4); also, in the neural processing results we verified that the performance on the test set significantly improves, thus allowing us to avoid the pre-processing previously necessary ([3]).

The spatial resolution on the focal plane (previously $2\mu m$ on each coordinate) is improved to $1\mu m$ in the along scan direction, and relaxed to $6\mu m$ in the across
scan direction (much less critical), compatibly with basic sampling requirements and computing time optimisation. Also, the computation is restricted to the region of interest corresponding to the planned readout windows (with margins).

The spectral representation has also been improved with respect to the previous monochromatic case. The polychromatic images are built by superposition of monochromatic PSFs, weighed with a distribution centred on the effective wavelengths 787 nm and 617 nm, estimated for blackbody sources at 3000 K (red) and 30000 K (blue), with FWHM of 200 nm, considered representative of the overall instrument and detector response. The basic parameters of the Gaia telescope (aperture, focal length) are used.

Notably, using the standard FFT approach, the sampling in either pupil or focal plane is variable with respect to wavelength, which is not physically sound. In our previous experiment, the focal plane sampling was retained, to avoid the need for image interpolation, thus inducing wavelength dependent sampling of the WFE corresponding to the current aberration case.

Realistic degradations due to nominal detector (finite pixel size and clocking in time-delayed integration) are taken into account, by filtering the PSF through representative functions. In the current approximation, the latter are rectangles with constant value over a width associated to the pixel size (10 \( \mu \)m) and the four-phase clocking (5 \( \mu \)m). Further realistic contributions (e.g. the modulation transfer function) can be easily introduced in future developments.

The polychromatic PSF for each source case is then integrated in the across scan direction (implementing the planned binning readout), and the one-dimensional moments are computed, for later neural processing. The chromaticity is estimated as COG difference between the blue and red star images.

A set of aberration cases is generated, in the regime of small image degradation, i.e. of reasonably good imaging performance. The Zernike aberration coefficients are generated from a Gaussian random distribution with \( \sigma = 50 \) nm for each component. The coefficient range is not configuration specific, but covers a set of mathematically possible cases, larger than the physically feasible optical systems. The corresponding RMS WFE on the aperture, averaged over the data set, is about 50 nm. The sample considered is thus representative of a range of realistic optical configurations.

Some of the moments do not have a significant trend with respect to chromaticity, and can therefore be neglected. The moment selection was verified on the NN using a pruning technique, i.e. selectively removing some of them and checking the convergence, until reaching the minimum number of parameters compatible with good training.

The NN inputs are thus defined in terms of the local instrument response, expressed by the local aberration values, then encoded in the moments for red and blue sources, which can be considered one as the reference and the other as the generic star of known spectral type. In particular, the COG of the reference object is the deviation of the image position with respect to an ideal system, and it is associated to the classical distortion; therefore, it is a system property which can be calibrated.
from the science data. The other inputs from the reference source are the image RMS width, the third and the fifth order moments. The inputs associated to the measured signal, from a star of different, known spectral type, are the third and fifth order moments.

The histogram of input chromaticity distribution in the test set is shown in Figure 5, and it is approximately Gaussian.

4 NEURAL PROCESSING AND RESULTS

Neural networks learn from examples; that is, given the training set of $N$ multi-dimensional data pairs $\{(x_i, F(x_i)) / x_i \in \mathbb{R}^P, F(x_i) \in \mathbb{R}^Q\}$, $i = 1, \ldots, N$, after the training if $x_i$ is the input to the network, the output is close to, or coincident with, the desired answer $F(x_i)$ and the network has generalization properties too, that is it gives as output $F(x_i)$ even if the input is only “close to” $x_i$, for instance a noisy or distorted or incomplete version of $x_i$; a comprehensive review on NN properties and applications can be found in [17].

In our work we use the multilayer perceptron, first introduced in 1986 (see [18]), as an extension of the perceptron model [19].
We use a sigmoidal NN with five inputs describing each blue and red star image plus the blue COG, one output (the chromaticity), and a single hidden layer with 30 units. The NN is trained by 10,000 iterations on a training set made by 10,000 instances, and its performance is verified on the 3,000 instances of the test set; in particular, the discrepancy between the NN output (estimated chromaticity) and target (actual chromaticity for the test set data instances) can be considered as the residual chromaticity after correction obtained from the NN results.

<table>
<thead>
<tr>
<th>Input Chromaticity (with outliers) [nm]</th>
<th>Residual Chromaticity (without outliers) [nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>−697.25</td>
</tr>
<tr>
<td>Mean</td>
<td>−1.20</td>
</tr>
<tr>
<td>Max.</td>
<td>750.62</td>
</tr>
<tr>
<td>RMS</td>
<td>190.50</td>
</tr>
</tbody>
</table>

Table 1. Main parameters of input and residual chromaticity in test set

The main statistical parameters of the residual chromaticity distributions, compared with the corresponding values of nominal chromaticity distribution (test set), are listed in Table 1. The central column refers to the diagnosis performances with outliers, i.e. on the whole test set; we noted, however, that on some values the so-called outliers (18 instances, 0.6% of data) performances are not so good, and they are outside the ±3σ interval. This problem will be the object of future works devoted to better understand this difficulty; however, we evaluated again the residual chromaticity distribution without these values obtaining the diagnosis performances listed in the right column; this distribution is shown in Figure 7. The 99.4% of originally processed data is, however, within the ±3σ interval.

Since the goal is the computation of output values coincident with the predefined target values, the plot of output vs. target, shown in Figure 8, should be ideally a straight line \(y = a + bx\) at angle \(\pi/4\), passing for the origin, i.e. with parameters \(a = 0, b = 1\).

We compute the best fit parameters of the NN output vs. target distribution and their standard deviation; the results, shown in Table 2, are quite consistent with the expectations.

<table>
<thead>
<tr>
<th>Offset (a)</th>
<th>(-2.498 ± 0.074)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope (b)</td>
<td>(0.998 ± 3.9e-004)</td>
</tr>
</tbody>
</table>

Table 2. Main parameters and errors of the linear fit

5 CONCLUSIONS

In this paper we use a neural network to diagnose and correct the systematic astrometric error of chromaticity, in a framework consistent with the mission Gaia
measurements. The science data are efficiently encoded in a set of low order image moments. The NN, with 300 internal nodes, is trained on a set of 8000 data instances, and evaluated on a test set of 2000 cases.

The NN diagnostics on the test set appears to be quite effective, as the residual chromaticity distribution, after data correction based on the NN results, is greatly reduced with respect to the initial distribution (Table 1).

Applying the NN output for correction of the chromaticity on the Gaia measurements, we may expect a significant reduction of the associated error; also, the residual chromaticity is expected to be random.

Future developments will include evaluation of the sensitivity to measurement noise, propagated through the image moments, which will induce practical limitations to the correction effectiveness, depending on the source brightness.

From the current results, NN diagnostics used to greatly reduce the chromatic error on astrometric measurements appears to be a tool able to produce effective results.


Rossella Cancelliere graduated with a Laurea in physics at the University of Torino (Italy), summa cum laude, in 1991. She was with several industries and academic research institutes, such as Polytechnic of Torino, Cselt (Torino), and University of Milano, where she took her Ph. D. in computational mathematics and operations research. Since October 1999 she is researcher at the University of Torino and since 2005 she is at the Department of Computer Science of the same university, where she works in the machine learning and data mining fields. Lecturer for the course of Neural Networks at the University of Torino since 2004.