The wavefront sensing making-of for THEMIS solar telescope

Michel Tallon, Éric Thiébaut, Maud Langlois, Bernard Gelly, Richard Douet, Clémentine Béchet, Loïc Denis, Gil Moretto

The wavefront sensing making-of for THEMIS solar telescope

Michel Tallon[a], Éric Thiébaut[a], Maud Langlois[a], Bernard Gelly[b], Richard Douet[b], Clémentine Béchet[c], Loïc Denis[d], and Gil Moretto[a]

[a]Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France
[b]CNRS UPS3718 - THEMIS, Via Lactea s/n, ES-38205 La Laguna, Canary Islands, Spain
[c]Institute of Mathematical and Computational Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile
[d]Univ. Lyon, UJM-Saint-Étienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Étienne, France

ABSTRACT

An adaptive optics system with a single deformable mirror is being implemented on the THEMIS 90cm solar telescope. This system is designed to operate in the visible and is required to be as robust as possible in order to deliver the best possible correction in any atmospheric conditions, even if wavefronts are sensed on some low-contrast solar granulation. In extreme conditions, the images given by the subapertures of the Shack-Hartmann wavefront sensor get randomly blurred in space, in the set of subapertures, and the distribution of blurred images is rapidly changing in time, some of them possibly fading away. The algorithms we have developed for such harsh conditions rely on inverse problem approach. As an example, with the gradients of the wavefronts, the wavefront sensor also estimates their errors, including their covariance. This information allows the control loop to promptly optimize itself to the fast varying conditions, both in space (wavefront reconstruction) and in time. A major constraint is to fit the calculations in a low-cost multi-core CPU. An overview of the algorithms in charge of implementing this strategy is presented, focusing on wavefront sensing.

Keywords: adaptive optics, wavefront sensing, centroiding method, registration method, solar telescope

1. INTRODUCTION

Adaptive optics (AO) is spreading among solar telescopes.[1] In this context, an adaptive optics system is being installed on THEMIS* 90 cm solar telescope at Teide Observatory, Tenerife, Canaria Islands. This new developpement offers the opportunity to implement new methods of reconstruction and control, driven by the requirement that the AO runs unsupervised at its best performance in any atmospheric conditions. The AO system is combined with a major refurbishment of the telescope optics, from the telescope secondary down to the spectrograph entrance, in order to keep the unique spectropolarimetric capabilities of this telescope.[2] These changes in the optical path should lead to an AO corrected telescope fit for high quality polarimetry.

As shown by figure 2, the AO system is based on a deformable mirror with 97 actuators (11×11), from ALPAO, and a Shack-Hartmann wavefront sensor with 76 subapertures (10×10) with a Fried geometry[3] (Fig. 3). In order to keep the cost as low as possible, the Real Time Computer is a single PC with 4 cores at 4 GHz, and run the control loop at 1 kHz.

This paper will focus on the new wavefront sensing method at the root of the control loop. The method aims at measuring optimally the displacements of the sub-images delivered by the wavefront sensor on the solar granulation seen with 2% contrast and a digitization on 8 bits only.

Since the day time atmospheric conditions may be quite instable, with sub-images randomly blurred in space and in time, some sub-images even fading away, the proposed method also estimate the errors corresponding

Send correspondence to Michel Tallon, E-mail: mtallon@obs.univ-lyon1.fr
*Télescope Héliographique pour l'Etude du Magnétisme et des Instabilités Solaires, *i.e.* Heliographic Telescope for the Study of Magnetism and Solar Instabilities

Figure 1. Shack-Hartmann wavefront sensor installed at THEMIS telescope.

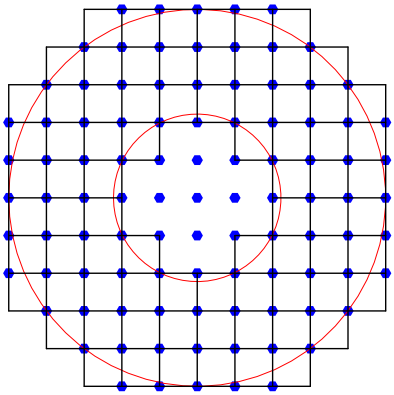

Figure 2. Geometry of the THEMIS adaptive optics system, with a 10×10 Shack-Hartmann wavefront sensor (76 active subapertures) and a ALPAO DM97 deformable mirror with 97 actuators. The red circles denote the shape of the telescope pupil with its central obscuration.

to the measurements obtained at each frame, including their covariances. These errors will allow wavefront reconstruction to optimally take into account the variations of the errors in the pupil and will allow the loop to optimize itself to fast varying conditions.

Since the measurements now come with their associated precision matrix (invert of the covariance matrix) in real time, the computation of the commands for the adaptive optics correction must also be able to take them into account. The usual method relies on a command matrix obtained with a pseudo inverse of the interaction matrix. This inversion is usually computed with truncated SVD and may include modal filtering.[4] Such an estimator may include fixed weights, for instance to take laser guide star elongation into account.[5] But here, this scheme would implies to compute a new command matrix at each frame since precision matrix changes at this rate. The solution is to compute the commands by an iterative method, without matrices,[6] since it only uses the direct model (sparse interaction matrix, weights, and priors) that can be changed on the fly. The algorithm to compute the commands from the wavefront sensor measurements is out of the scope of this paper where we focus on wavefront sensing on open-loop real data acquired with THEMIS wavefront sensor.

The wavefront sensing method presented in the following is based on the fit of a reference image used as a model, on the pixels of the sub-images. The covariance matrix of the measurements can be obtained if the error on the pixel values is known. So the method starts with a specific calibration of the detector in order to estimate the error of each pixel value.

The next sections will overview successively the pre-processing of the pixels, with a particular step to compute an estimate of the error on each pixel value, the estimation of the slopes and their covariances from the pre-processed pixels, and the estimation and the update of the reference image used as a model for the estimation of the slopes.

2. PRE-PROCESSING OF RAW PIXELS

Pixel values are at the root of the whole processing. In order to estimate the covariance of the errors on the slope measurements, we need to estimate the variances of the pixel values. This implies an additional step to the usual dark / flat correction process, which is presented in this section. For each pixel i, we seek for the incident flux in arbitrary unit and the standard deviation of the errors on the estimated flux. In order to simplify the derivation of the equations, we assume here that all the necessary calibration frames (see after) are available with the same

exposure time as the data (wavefront sensor frames). This assumption is quite practicable here since the flux from the Sun is fairly stable. With this assumption, a pixel raw value in ADU, r_i, can be written as:

$$r_i = \frac{t_i\,\phi_i}{g_i} + b_i + n_i, \tag{1}$$

where t_i is the transmission from the sky down to the pixel i, including the quatum efficiency of the detector, ϕ_i is the flux of the source in an arbitrary unit proportional to photons, g_i is the gain (e^-/ADU), and b_i and n_i are respectively the bias and the noise, both in ADU.

We use tree types of calibration frames obtained with the same exposure time:

- *dark frames*: $\phi_i = 0$, *i.e.* the pixel does not receive any flux. A large set of dark frames are recorded to compute the mean and variance on each pixel, denoted $\mathbb{E}(r_i^{\mathsf{dark}})$ and $\mathrm{Var}(r_i^{\mathsf{dark}})$ respectively.
- *flat frames*: $\mathbb{E}(\phi_i) = \mathbb{E}(\phi^{\mathsf{flat}})$, *i.e.* we assume that all the pixels receive the same flux on average, even if this flux varies in time. This flat frame is obtained on the Sun while the telescope is speedily wandering other the solar granulation. We only need the mean of a large set of frames, denoted $\mathbb{E}(r_i^{\mathsf{flat}})$.
- *static frames*: $\mathrm{Var}(t_i\,\phi_i) = \mathbb{E}(t_i\,\phi_i)$, *i.e.* we can assume Poisson statistics of the counts, which holds even if the pixels receive different amounts of light. In this case, we use an internal stabilized source that shed some, non-uniform but stable in time, light on the detector (a Mikrotron EoSens 4CXP camera). Here again, a large set of frames is used to compute the empirical mean and variance on each pixel, denoted $\mathbb{E}(r_i^{\mathsf{stat}})$ and $\mathrm{Var}(r_i^{\mathsf{stat}})$ respectively.

Compared to the usual detector calibration, we thus need the variance of the *dark frames*, and a set of so-called *static frames*. In principle, a source could deliver both *flat* and *static* frames, but the "flattening method" generally introduces fluctuations of the source, as it is the case with solar granulation.

The *flat frames* are obtained through the field-stop of the wavefront sensor; so their mean is also used to determine the regions of interest (ROI) on the detector for each sub-image. The flat correction also account for the non-uniform transmission of the optics (*e.g.* vignetting by the microlenses).

The calibrated value of the pixel i is obtained from a linear combination of raw value r_i, as usual:[7]

$$d_i = \alpha_i(r_i - \beta_i). \tag{2}$$

By assuming $\mathbb{E}(d_i) = \mathbb{E}(\phi_i)/\mathbb{E}(\phi^{\mathsf{flat}})$, we get:

$$\alpha_i = \frac{1}{\mathbb{E}(r_i^{\mathsf{flat}}) - \mathbb{E}(r_i^{\mathsf{dark}})}, \tag{3}$$

$$\beta_i = \mathbb{E}(r_i^{\mathsf{dark}}). \tag{4}$$

Thus if the flat frame is obtained on the Sun itself, the calibrated values of the pixels are around unity.

Assuming that the value d_i is an estimation of the expected value (we have a single sample for one pixel!), the variance of the pixel value is also obtained from a linear combination:

$$\mathrm{Var}(d_i) \approx \frac{\max(d_i, 0) + v_i}{u_i}, \tag{5}$$

with

$$g_i = \frac{\mathbb{E}(r_i^{\mathsf{stat}}) - \mathbb{E}(r_i^{\mathsf{dark}})}{\mathrm{Var}(r_i^{\mathsf{stat}}) - \mathrm{Var}(r_i^{\mathsf{dark}})}, \tag{6}$$

$$u_i = g_i\left(\mathbb{E}(r_i^{\mathsf{flat}}) - \mathbb{E}(r_i^{\mathsf{dark}})\right), \tag{7}$$

$$v_i = g_i\,\frac{\mathrm{Var}(r_i^{\mathsf{dark}})}{\mathbb{E}(r_i^{\mathsf{flat}}) - \mathbb{E}(r_i^{\mathsf{dark}})}. \tag{8}$$

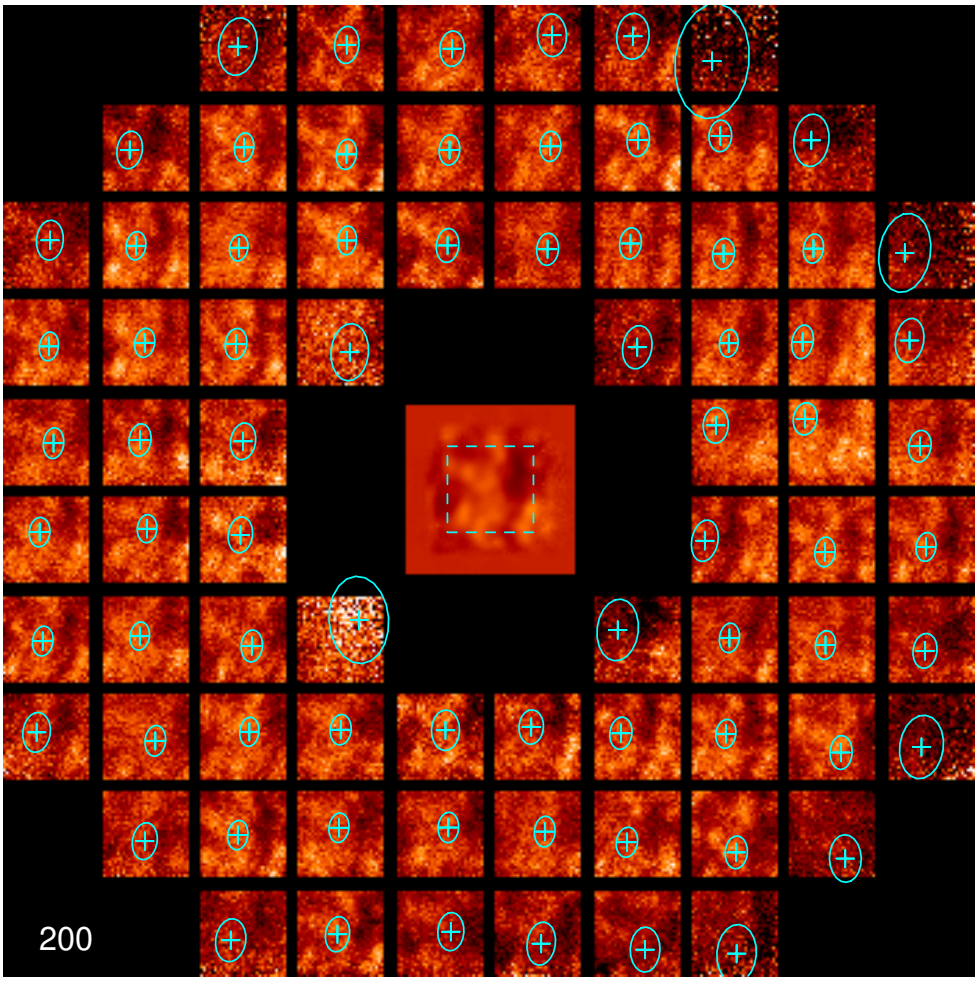

Figure 3. Superimposition of a wavefront frame pre-processed as explained in Sec. 2, and the result of the measurement process as explained in Sec. 3. The 76 sub-images show a 10″ field-of-view on 30 × 30 pixels, obtained with a ∼ 200µs exposure time at 1kHz. The contrast of the granulation seen in each subaperture is ∼ 2%. The crosses show the estimated locations of the sub-images for this frame, and the ellipses the estimated standard deviation of these locations. This frame shows that the accuracy along the x-axis is better than the one along the y-axis, and that subapertures at the edges of the pupil are less accurate. The reference image shown at the center is estimated from the previous frames on a larger field-of-view (dashed square) and with more details than the field-of-view of the subapertures (*cf.* Sec. 4).

The estimation of the errors on the pixels thus needs the additional computation of $\mathrm{Var}(r_i^{\mathsf{dark}})$, $\mathbb{E}(r_i^{\mathsf{stat}})$ and $\mathrm{Var}(r_i^{\mathsf{stat}})$.

In the following, we gather all the pixels d_i of a sub-image k in a single vector $\boldsymbol{d}_k$. In the same way, their corresponding weights $1/\mathrm{Var}(d_i)$ are gathered in the precision matrix $\mathbf{W}_k$ of sub-image k. Since the covariance terms are neglected, the matrix $\mathbf{W}_k$ is diagonal.

3. MEASUREMENT WITH A KNOWN REFERENCE IMAGE

Assuming Gaussian noise, maximum likelihood amounts to minimizing the following objective function:

$$\psi(\boldsymbol{\theta}) = \sum_{k=1}^{n} \left\{ \eta_k \, \|\boldsymbol{d}_k - \alpha_k \, \mathbf{R}_k(\boldsymbol{s}_k).\boldsymbol{r}\|^2_{\mathbf{W}_k} - m_k \, \log \eta_k \right\} + \mu \, \|\mathbf{D}.\boldsymbol{r}\|_2^2 \tag{9}$$

with $\boldsymbol{d}_k$ the calibrated pixel values in sub-image k, $\mathbf{W}_k$ the corresponding precision matrix, m_k the number of pixels in the sub-image k, $\mathbf{R}_k(\boldsymbol{s}_k)$ an operator that shifts the *reference image* $\boldsymbol{r}$ by an offset $\boldsymbol{s}_k$, α_k a rescaling of

the reference image for modeling the scintillation and η_k a rescaling of $\mathbf{W}_k$ to take modeling errors into account. Thus we aim here at fitting each sub-image $\boldsymbol{d}_k$ with a known *reference image*, $\boldsymbol{r}$, that can be shifted by an offset $\boldsymbol{s}_k$ and rescaled by a factor α_k. Because of registration errors, we expect that sub-images do not perfectly match the reference image, for that reason we allow the precision matrix to be rescaled by a quality factor η_k for each sub-image. As the quality factors are part of the parameters to fit, the associated logarithm term must be kept in the objective function.

In Eq. (9), the term $\mu \|\mathbf{D}.\boldsymbol{r}\|_2^2$ is meant to regularize the estimation of the reference image $\boldsymbol{r}$ that will be addressed in the next section. In this section, we assume the reference image $\boldsymbol{r}$ fixed and known, this term has no effect.

We are only looking for the offsets $\boldsymbol{s}_k$ and their covariances, but need to fit in real time all the parameters $\boldsymbol{\theta} = \{\eta_k, \alpha_k, \boldsymbol{s}_k\}_{k=1,\dots,n} \cup \boldsymbol{r}$.

The operations applied to the reference image are non linear with respect to the shifts $\boldsymbol{s}_k$. A linearization of $\alpha_k \mathbf{R}_k(\boldsymbol{s}_k).\boldsymbol{r}$ is necessary to derive an algorithm fast enough to be applied in real time. This linearization step relies on a first order expansion and a change of variable around an expected value of $\boldsymbol{s}_k$, as detailed in Thiébaut *et al.*[8] The likelihood term for one sub-image now writes:

$$\varphi_k(\boldsymbol{u}_k) = \|\boldsymbol{d}_k - \mathbf{H}_k.\boldsymbol{u}_k\|^2_{\mathbf{W}_k} = \gamma_k - 2\,\boldsymbol{b}_k^{\mathrm{T}}\boldsymbol{u}_k + \boldsymbol{u}_k^{\mathrm{T}}\mathbf{A}_k.\boldsymbol{u}_k, \tag{10}$$

with

$$\boldsymbol{u}_k = \alpha_k \begin{pmatrix} 1 \\ x_k \\ y_k \end{pmatrix} \tag{11}$$

where (x_k, y_k) is the sought relative displacement from the predicted position, $\mathbf{H}_k$ is a $m_k \times 3$ matrix obtained from a first order expansion[8] of the shifted reference image, and:

$$\gamma_k = \|\boldsymbol{d}_k\|^2_{\mathbf{W}_k}, \tag{12}$$

$$\boldsymbol{b_k} = \mathbf{H}_k^{\mathrm{T}}.\mathbf{W}_k.\boldsymbol{d}_k, \tag{13}$$

$$\mathbf{A}_k = \mathbf{H}_k^{\mathrm{T}}.\mathbf{W}_k.\mathbf{H}_k. \tag{14}$$

The solution of Eq. (10) is:

$$\widehat{\boldsymbol{u}}_k = \mathbf{A}_k^{-1}.\boldsymbol{b}_k, \tag{15}$$

and the corresponding value of $\varphi_k(\boldsymbol{u}_k)$ at the minimum is:

$$\varphi_k(\widehat{\boldsymbol{u}}_k) = \gamma_k - \boldsymbol{b}_k^{\mathrm{T}}.\widehat{\boldsymbol{u}}_k \tag{16}$$

The next step is to estimate the rescaling of the weights $\mathbf{W}_k$ from Eq. (9). After this linearization and solving for $\widehat{\boldsymbol{u}}_k$ for each sub-image k, Eq. (9) now writes:

$$\psi(\boldsymbol{\theta}) = \sum_{k=1}^{n} \{\eta_k\, \varphi_k(\widehat{\boldsymbol{u}}_k) - m_k\, \log \eta_k\}, \tag{17}$$

the value of the likelihood term, $\varphi_k(\widehat{\boldsymbol{u}}_k)$, being known from Eq. (16). Minimizing $\psi(\boldsymbol{\theta})$ for each sub-image, yields:

$$\eta_k = \frac{m_k}{\varphi_k(\widehat{\boldsymbol{u}}_k)}. \tag{18}$$

For each sub-image k, the covariance matrix, $\mathbf{C}_k$, of the parameters (α_k, x_k, y_k) is approximately given by:[8]

$$\mathbf{C}_k \approx \frac{1}{\eta_k}\, \mathbf{J}_k^{\mathrm{T}}.\mathbf{A}_k^{-1}.\mathbf{J}_k, \tag{19}$$

with $\mathbf{J}_k$ the Jacobian matrix of the partial derivatives of the non-linear relationship between $\boldsymbol{u}_k$ and (α_k, x_k, y_k).

Figure 3 shows an example of a processed frame. On top of a frame from the wavefront sensor, pre-processed as detailed in Sec. 2, the crosses show the estimated locations (x_k, y_k) of the sub-images, and the ellipses denote the estimated standard deviation around these locations, from the coefficients of $\mathbf{C}_k$. The ellipses have been scaled so that the size of a sub-image correspond to the size of a pixel, so that the standard deviation of the position is ± 0.5 pixel when the diameter of the an ellipse equals the size of a sub-image. We can see variations of the errors across the pupil, with larger errors from the subapertures at the edge of the pupil. The ellipse are also elongated because of the particular structure of the reference image shown in the middle of the image: the errors are not the same along x and y axis, with some slight correlation between the coordinates.

4. UPDATE OF THE REFERENCE IMAGE

In the previous section, by assuming the reference image $\boldsymbol{r}$ is known, the optimization of the general cost fonction $\psi(\boldsymbol{\theta})$ given by Eq. (9) allowed all the parameters but $\boldsymbol{r}$ to be determined in the set of the parameters $\boldsymbol{\theta} = \{\eta_k, \alpha_k, \boldsymbol{s}_k\}_{k=1,\dots,n} \cup \boldsymbol{r}$. In an alternate approach, we now look for the reference image $\boldsymbol{r}$ itself, assuming all the other parameters are known.

Denoting $\mathbf{G}_k = \alpha_k \mathbf{R}_k(\boldsymbol{s}_k)$ (*i.e.* shifting the reference image by an offset $\boldsymbol{s}_k$, and rescaling by factor α_k), we can write the significant terms of Eq. (9) for our purpose as:

$$\psi'(\boldsymbol{\theta}) = \sum_{k=1}^{n} \|\boldsymbol{d}_k - \mathbf{G}_k.\boldsymbol{r}\|^2_{\eta_k \mathbf{W}_k} + \mu \|\mathbf{D}.\boldsymbol{r}\|_2^2 . \tag{20}$$

Thus the expression is reduced to regularized least-squares. The underlying approximation here for a fast computation when applying $\mathbf{G}_k$ is to shift $\boldsymbol{r}$ by $\boldsymbol{s}_k$ rounded at the nearest pixel, thus avoiding the need of any interpolation.

The regularization term is mandatory here because the reference image is used as the model to be fitted and need to be known everywhere it is needed, with a good accuracy. So the estimation of the reference image must be extrapolated on all the field-of-view, enlarged by the dynamic of the wavefront sensor. We assume that the enlarged field-of-view is twice the field-of-view of the sub-images. The chosen priors, $\mathbf{D}$, is a difference operator that computes finite differences of the reference image.

Minimizing Eq. (20) for $\boldsymbol{r}$ yields:

$$\left(\sum_{k=1}^{n} \eta_k \, \mathbf{G}_k^{\mathrm{T}}.\mathbf{W}_k.\mathbf{G}_k + \mu \, \mathbf{D}^{\mathrm{T}}.\mathbf{D} \right) \boldsymbol{r} = \sum_{k=1}^{n} \eta_k \, \mathbf{G}_k^{\mathrm{T}}.\mathbf{W}_k.\boldsymbol{d}_k . \tag{21}$$

This equation is solved for $\boldsymbol{r}$ by using conjugate gradient method. We can notice that the right hand side expression is the weighted sum of the recentered sub-images.

In this process, the chicken or the egg problem is solved at the first frame by bootstrapping the method with setting $\boldsymbol{r}$ as the simple weighted average of all the sub-images. The first set of parameters are determined with this first reference image, up to the computation of a new reference image by solving Eq. (21). This new reference image is used for the next frame, and the first one is deleted.

For all the following frames, a new reference image is computed as soon as the measurements are sent to the controller that computes the commands. This new reference is merged with the previous one by using a leaking integrator.

By merging the recentered sub-images spatially in an optimal way, the reference image gets to be known in a field larger than the field-of-view of the sub-images. This allows all the field-of-view of the sub-images to be used to get the measurements of their shifts. An example of such an enlargement is shown on Fig. 3: the reference image is shown at the center of the image. This reference image is much less noisy than one sub-image chosen as a reference. Furthermore, the reference image is fixed and just evolves with the granulation pattern, thus allowing tip/tilt to be continuously measured while maintaining the same line of sight for a long time.

5. CONCLUSION

The presented wavefront sensing method for solar adaptive optics delivers for each frame, both the slope measurements and their covariance matrix. By using this information, the control loop can be robust to the fast spatial and temporal variations of the measurement errors. The quality of the measurements is improved by using an estimate of the reference image optimally built from the merging of all the sub-images, thus known with more details on a larger field-of-view and with much less noise. Furthermore, the slope measurements obtained by fitting such a reference image on the sub-images has been shown to be optimal.[8] The bootstrapping of the algorithm has been checked on open-loop data with various samples of the granulation.

At this time, the method is being implemented in the Real Time Computer in order to close the loop at 1 kHz. This goal is yet to be demonstrated on sky.

ACKNOWLEDGMENTS

This project has been co-funded by the European Commission's FP7 Capacities Programme under Grant Agreement number 312495, and the Centre National de la Recherche Scientifique.

REFERENCES

[1] Rimmele, T. R. and Marino, J., "Solar Adaptive Optics," *Living Reviews in Solar Physics* **8**, 2 (June 2011).

[2] Gelly, B., Langlois, M., Moretto, G., Douet, R., Lopez Ariste, A., Tallon, M., Thiébaut, E., Geyskens, N., Lorgeoux, G., Léger, J., and Le Men, C., "New life for the THEMIS solar telescope," in [*Ground-based and Airborne Telescopes VI*], *Proc. SPIE* **9906**, 99065A–99065A–11 (2016).

[3] Fried, D. L., "Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Am.* **67**(3), 370–375 (1977).

[4] Gendron, E. and Léna, P., "Astronomical adaptive optics. i. modal control optimization," *A&A* **291**, 337–347 (1994).

[5] Tallon, M., Tallon-Bosc, I., Béchet, C., and Thiébaut, E., "Wavefront reconstruction with elongated spots," in [*ELT-DS WP09900 teleconference*], (February 11 2008). Lyon.

[6] Thiébaut, E. and Tallon, M., "Fast minimum variance wavefront reconstruction for extremely large telescopes," *J. Opt. Soc. Am. A* **27**, 1046–1059 (Apr. 2010).

[7] Léna, P., Rouan, D., Lebrun, F., Mignard, F., and Pelat, D., [*Observational Astrophysics*], Astronomy and Astrophysics Library, Springer-Verlag Berlin Heidelberg, 3rd ed. (2012).

[8] Thiébaut, É., Tallon, M., Denis, L., Langlois, M., Béchet, C., Moretto, G., and Gelly, B., "Innovative real-time processing for solar adaptive optics," in [*Adaptive Optics Systems VI*], *SPIE* **10703**, 107031I (July 2018).