A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
03 (03.13.6; 03.13.4; 08.22.3; 12.07.1)

ASTRONOMY AND ASTROPHYSICS
March 8, 2000

Image subtraction using a space-varying kernel.

C. Alard[1,2]

[1] DASGAL, 61 Avenue de l'observatoire, F-75014 Paris, France
[2] Institut d'Astrophysique de Paris, 98bis Boulevard Arago, F-75014

Received ......; accepted ......

Abstract. Image subtraction is a method by which one image is matched against another by using a convolution kernel, so that they can be differenced to detect and measure variable objects. It has been demonstrated that constant optimal-kernel solutions can be derived over small sub-areas of dense stellar fields. Here we generalize the theory to the case of space-varying kernels. In particular, it is shown that the CPU cost required for this new extension of the method is almost the same as for fitting a constant kernel solution. It is also shown that constant flux scaling between the images (constant kernel integral) can be imposed in a simple way. The method is demonstrated with a series of Monte-Carlo images. Differential PSF variations and differential rotation between the images are simulated. It is shown that the new method is able to achieve optimal results even in these difficult cases, thereby automatically correcting for these common instrumental problems. It is also demonstrated that the method does not suffer due to problems associated with under-sampling of the images. Finally, the method is applied to images taken by the OGLE II collaboration. It is proved that, in comparison to the constant-kernel method, much larger sub-areas of the images can be used for the fit, while still maintaining the same accuracy in the subtracted image. This result is especially important in case of variables located in low density fields, like the Huchra lens. Many other useful applications of the method are possible for major astrophysical problems; Supernova searches and Cepheids surveys in other galaxies, to mention but two. Many other applications will certainly show-up, since variability searches are a major issue in astronomy.

Key words: Methods: numerical – Methods: statistical – Stars: variables: general – Cosmology: gravitational lensing

1. INTRODUCTION

The way variability is analyzed in astronomical images has been deeply influenced by the development of the microlensing experiments (EROS, OGLE, MACHO, DUO). The enormous quantity of CCD images produced by these experiments has motivated the development of fast and accurate techniques to analyze variability. The first implementations of variability search were based on an analysis of catalogues of stellar objects. The production of the catalogues has been almost exclusively done using the DoPHOT software (Mateo & Schechter). This type of analysis is still used in microlensing experiments, and has proved very successful.

However since we are concerned only with differential photometry, such data processing is certainly not optimal. Purely differential methods, like image subtraction, would appear to be better suited for study of variability in astronomical images. However the development of image subtraction techniques has been long delayed by the inherent difficulty and complexity of the method.

The first successful implementation of the method was by Tomaney & Crotts (1996). Further progress has been made by the creation of the fast Optimal Image Subtraction (OIS) method (Alard & Lupton 1998). Due to its ability to solve the full least-square problem, the OIS method proved to produce subtracted images and light curves with an accuracy approaching closely the photon noise limit (Alard 1999). The OIS method is very efficient for crowded stellar fields, like those encountered in microlensing experiments, and could readily be used for massive processing of microlensing data. The current implementation of the OIS works by dividing the field into small sub-areas where constant kernel solutions are derived. In dense crowded fields, sub-areas as small as 128×128 pixels can be used. At this scale kernel variations can usually be ignored to a good approximation. However in case of very bad optics or less dense fields, the constant kernel approximation does not hold any more. For instance, in high latitude fields taken for supernovae search, there are often insufficient bright objects per unit area over which the kernel variations can be ignored. We are then forced to try to make a self consistent fit of the kernel variations, as described in Alard & Lupton (1998). It is important to notice though that even if the spatial variations are fitted to order 1 only, the cost to build the least-square matrix will be about 9 times larger than for a constant kernel solution. Some situations might require order 2 or 3 or even more. Order 3 requires roughly 100 times more calculations than a constant kernel solution. Clearly the problem quickly becomes

Send offprint requests to: C. Alard

intractable, and that one of the main advantages of the OIS, the fast computing time, will be completely lost.

Fortunately, it will be demonstrated in this article that the fit of the spatial variations of the kernel can be achieved for little additional computing cost, given a reanalysis of the problem.

2. Basic equations.

Here we briefly summarize the basics of image subtraction (Alard & Lupton 1998). The essence of the method is to find a convolution kernel (K), that will transform a reference image (R) to fit a given image I.
In terms of least-squares this is equivalent to solving for a kernel that will minimize the sum:

$$\sum_i \left([R \otimes K](x_i, y_i) - I(x_i, y_i)\right)^2 . \tag{1}$$

Provided the kernel can be decomposed onto basis functions, the above equations become a simple linear least-squares problem. For the kernel decomposition we take:

$$K(u,v) = \sum_n a_n \; K_n(u,v)$$

with:

$$K_n(u,v) = e^{-(u^2+v^2)/2{\sigma_k}^2} \; u^i \; v^j$$

And the generalized index $n = (i, j, k)$

The kernel solution can be calculated by solving the following linear system:

$$M\mathbf{a} = \mathbf{B} \tag{2}$$

With:

$$M_{ij} = \int [R \otimes K_i](x,y) \; \frac{[R \otimes K_j](x,y)}{\sigma(x,y)^2} \; dxdy$$

And:

$$B_i = \int I(x,y) \; \frac{[R \otimes K_i](x,y)}{\sigma(x,y)^2} \; dxdy$$

3. Solving for space-varying kernel solution with minimum computing time.

Most of the computing time involved in the computation of the kernel solution is spent in the calculation of the least-square matrix. The solution of the linear system itself takes an almost negligible amount of time. The computing time to build the matrix goes like the square of the number of coefficients. Thus the problem is that if we introduce new coefficients in order to fit the kernel variations, the cost of the calculation will quickly become prohibitive. To be more specific, let's derive analytical formulae for the spatial variations of the kernel. At each position (x,y) we can develop the kernel on the basis functions K_n. Thus the coefficients a_n will be functions of (x,y). Consequently the kernel can be written:

$$K(u,v) = \sum_n a_n(x,y) \; K_n(u,v)$$

We assume that the coefficients a_n are smooth function of x and y. For simplicity we adopt a polynomial function of degree d_1.

$$a_n(x,y) = \sum_{i,j} b_{i,j} \; x^i \; y^j$$

The relevant least-squares vectors are:

$$V_n(x,y) = \; W_m(x,y) \times P_q(x,y)$$

With:

$$W_m(x,y) = \frac{[R \otimes K_m](x,y)}{\sigma(x,y)}$$

$$P_q(x,y) = x^i \; y^j$$

And:

$$q = i + j \times d_1$$

$$nc_1 = \frac{(d_1+1)(d_1+2)}{2}$$

$$n = q \; + \; m \times nc_1$$

The elements of the least-squares matrix can be expressed as scalar products of the least-squares vectors:

$$M_{n1,n2} = \int W_{m1}(x,y) \; W_{m2}(x,y) \; P_{q1}(x,y) \; P_{q2}(x,y) \; dxdy$$

The above integral extends over the entire image. However, we can always divide the integration domain into small rectangular sub-areas. Within each small area, one can assume that the kernel is constant. It corresponds to approximating x and y by the coordinates of the centers of the sub-areas, $(x_k, \, y_k)$.
The matrix elements can then be written:

$$M_{n1,n2} = \sum_k Q^k_{m1,m2}(x,y) \; P_{q1}(x_k,y_k) \; P_{q2}(x_k,y_k) \tag{3}$$

with the following definition for the integral in the rectangular domain D_k:

$$Q^k_{m1,m2} = \int_{D_k} W_{m1}(x,y) \; W_{m2}(x,y) \; dxdy.$$

A similar method can be applied to expand the vector on the right hand side of eq. 2.

$$B_n = \sum_k \; W^k_m(x,y) P_{q1}(x_k,y_k) \tag{4}$$

With:

$$W^k_m(x,y) = \int_{D_k} I(x,y) \; \frac{W_{m1}(x,y)}{\sigma(x,y)} \; dxdy$$

We see that the matrix elements corresponding to fitting the spatial variations of the kernel can be deduced from the matrix elements $Q^k_{m1,m2}$ for only the cost of a summation over the number of sub-areas. The matrix $Q^k_{m1,m2}$, corresponds to the constant kernel solution, and has a much smaller number of elements than the matrix $M^k_{n1,n2}$. Once $Q^k_{m1,m2}$ has been estimated, $M_{n1,n2}$ can be calculated using the expansion given in Eq. 3. This procedure results in a drastic economy in computing time, since the summations inside the individual sub-areas are avoided for the calculation of the matrix elements. Using Eq. 4 the same type of expansion applies also to the calculation of the vector B_n. Since the individual sub-areas should be slightly larger than the kernel, a typical number for the surface of a sub-area is 30×30 pixels. It means that in computing application we make a saving of about a factor 1000 in the calculation of the matrix elements of the matrix $M_{n1,n2}$ which corresponds to the spatial variations. In practice, fitting the kernel variations costs only about 20 % more than making a constant kernel fit.

In high sparse fields, e.g. at high latitudes, the only useful areas are those centered on bright (high S/N) objects. It usually results in a drastic reduction in the number of pixels to fit. Even in crowded fields, most of the information is contained in some high density regions. Thus in practice the fit can be restricted to a fraction of the image.

4. Imposing constant flux scaling.

In practical applications, we expect the flux scaling factor between the 2 images being matched to be constant. As a result of the large number of degrees of freedom, the flux scaling factor, which is the sum of the local kernel, in general will not be constant. This will lead to imperfect flux conservation. Since flux conservation is extremely important for photometry, it is essential to add this constraint to the fit.

Least squares fitting with constraints is not always easy to implement. In this case the constraint will be that the sum of the kernel be constant at all points of the image. By using the derivative of the sum one may solve this problem by the technique of the Lagrange multipliers. However, in this case the number of Lagrange multipliers will be close to the number of basis vectors used to model the spatial variations. Thus it would require solving a large number of additional equations, and will not be very efficient.

Fortunately, there is a more elegant way to solve this constraint. One can always consider using a new basis of vectors for the kernel that are linear combinations of the previous vectors. We can construct the new set in such a way that all basis vectors, except the first one, have zero sums. Some of the vectors will already have zero sums, and we will keep them unchanged. For the others, except the first one, we make the combination:

$$K_n^{'} = K_n - K_0.$$

Provided the vector K_n are normalized, the vectors $K_n^{'}$ will all have zero sums except $K_0^{'}$ which will be equal to K_0.
Then all we have to do is expand in image coordinates all the coefficients of the kernel expression, except the first coefficient. It will result in the following kernel formula:

$$K(u,v) = a_0\ K_0(u,v) + \sum_{1,N} a_n(x,y)\ K_n^{'}(u,v)$$

One can easily check that such a kernel has a constant sum, since all vectors whose coefficients depend upon x and y have zero sums.

Of course, this method can be used to calculate, with a minimum number of operations, the global matrix from the zero order matrix. In practice this method gives very good numerical conservation of the kernel sum, this is an important asset making the method more robust and more accurate.

5. Numerical simulations.

5.1. Test with variable PSF.

5.1.1. The simulated images.

To check the ability of the method to reconstruct the spatial variations of the kernel, a series of tests were performed with simulated images. The images are generated by putting stars randomly in the image, with a magnitude distribution corresponding to a bulge luminosity function. Noise is added in the images according to Poisson statistics. For the reference image we take a constant gaussian PSF:

$$\Phi_0(u,v) = e^{-\alpha(u^2+v^2)}$$

In our simulation we take $\alpha = 0.5$.

For the other image we simulate a variable PSF by taking the sum of 2 gaussians with different widths and a relative weighting which is a function of the position in the image. We also introduce a normalization factor for conservation of the flux.

$$\Phi(u,v) = \frac{e^{-\alpha(u^2+v^2)} + r(x,y) \times e^{-\frac{\alpha(u^2+v^2)}{4}}}{1 + 4\, r(x,y)}$$

For the spatial variation we take:

$$r(x,y) = \frac{0.25\, y}{w}$$

Where w is the size of the simulated image.
The resulting images are presented in Fig. 1.

5.1.2. Image subtraction.

First, a subtracted image with constant kernel solution was constructed. We used a set of parameters for the kernel similar to Alard & Lupton (1998). For the fit with variable kernel solution we used a polynomial of order 2, since the variations of the coefficients of the PSF function are nearly parabolic. The resulting subtracted images are presented in Fig. 2.

5.1.3. Noise estimation.

Noise in the subtracted image has 2 origins, the noise in the image to fit, and the noise in the reference image convolved with the kernel solution .

In our images the Poisson noise can be extremely well approximated locally by a gaussian distribution with $\sigma = \sqrt{N}$.

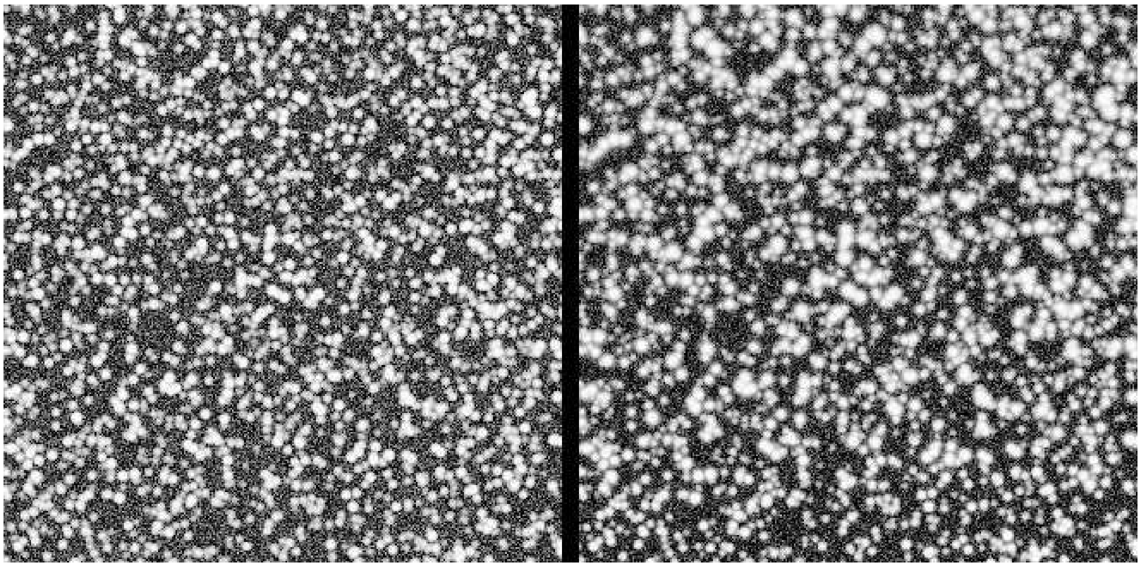

Fig. 1. Simulated of crowded field images. On the left is the image with constant PSF, and on the right is the image with PSF variations along the Y axis. Note the large amplitude of the PSF variations. A total of 2500 stars has been included in this simulation.

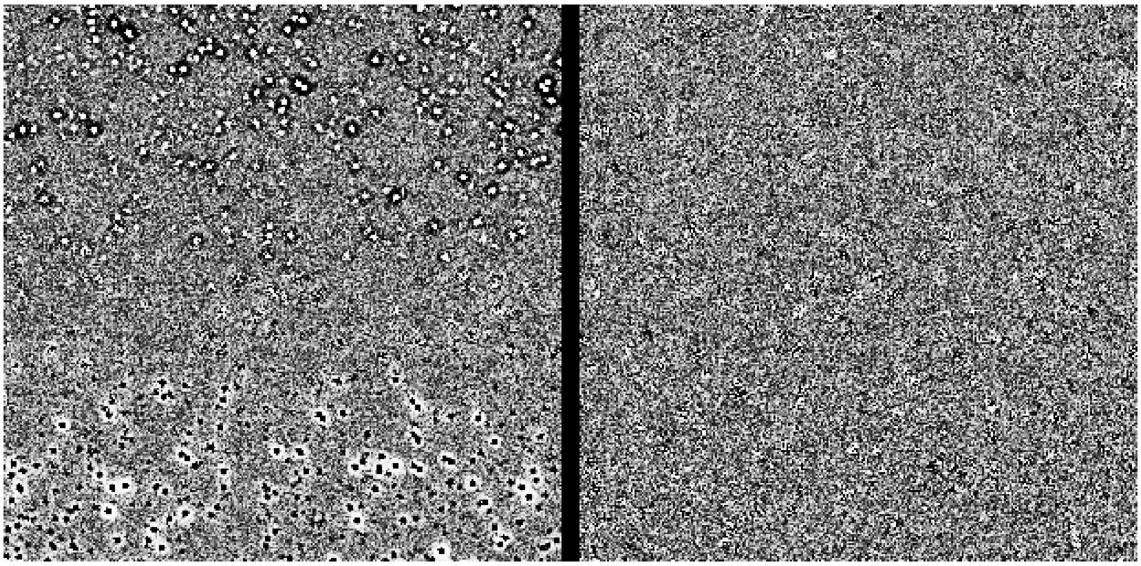

Fig. 2. on the left is the subtracted image obtained with constant kernel solution. Note the systematic pattern along the Y axis due to the kernel variations. On the right we present the subtracted image obtained by fitting the spatial variations of the kernel to order 2.

However this is not true for the reference image, since it has been transformed by convolution. The noise in the convolved image can nonetheless be estimated in a straightforward way. The convolution will result in the combination of different Gaussian distributions, with different σs and weights. Using the same notation where I is the image to fit and R is the reference image, we now define the IC to be the reference image convolved with the kernel solution,

$$IC_i = \sum_j R_{i-j}\, K_j.$$

The combination of the two gaussian distributions will result in a gaussian distribution, and the resulting σ of the distribution can be estimated by calculating the variance:

$$\sigma_i^2 = \sum_j var(R_{i-j})\, K_j^2 = \sum_j R_{i-j}\, K_j^2$$

Thus we see that the local variance of the image can be estimated by convolving the image with a filter that is just the square of the initial convolution filter. Consequently we define a *Poisson deviation* as:

$$\delta = \sqrt{\sigma_i^2 + I_i}$$

The histograms of the pixels in the subtracted images normalized by the Poisson deviation are presented in Fig. 3. It is interesting to note that order 2 is sufficient to produce a χ^2 per degree of freedom (χ^2/Dof) which is extremely close to 1.

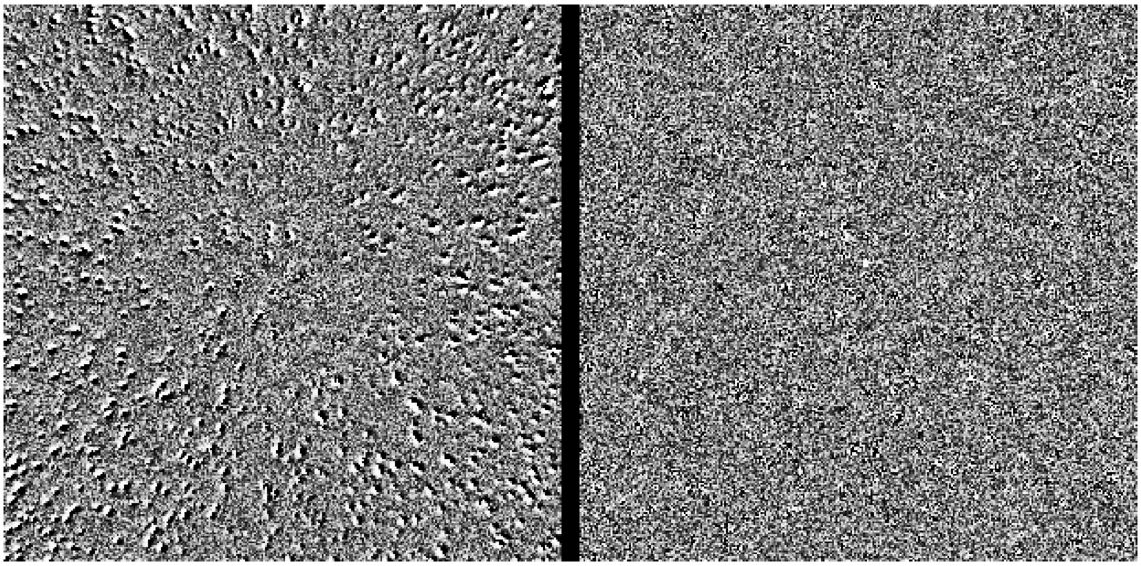

Fig. 4. Left is the subtracted image obtained with constant kernel solution. Note the systematic pattern due to the differential rotation between the images.Next is the subtracted image obtained by fitting the spatial variations of the kernel to order 2. Note the complete disappearance of the systematic patterns.

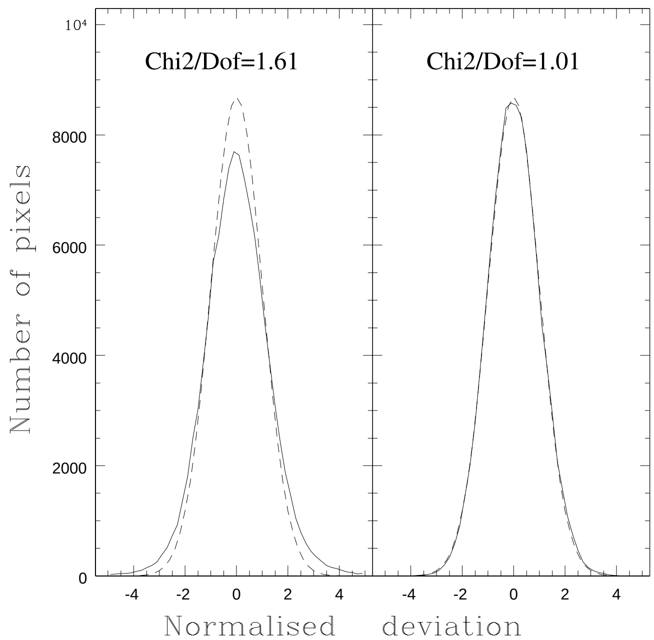

Fig. 3. Histograms of the normalized deviations in the subtracted images presented in Fig. 2. Left is the histogram for constant kernel solution, and right is the histogram for a fit of kernel variation to order 2. The pixels in the subtracted images have been normalized by the Poisson deviations (see text for details). The dashed curve is gaussian with $\sigma = 1$.

5.2. Checking the ability of the method to correct the astrometric registration.

Image registration is performed by calculating a polynomial transform from the positions of bright objects. However, there is no guarantee that this procedure is optimal. In case the image registration to the reference frame is not perfect, a simple translation can be taken into account with a constant kernel solution. However more complex features like differential rotation, or differential distortion between the images cannot be corrected with constant kernel solution. But they can be corrected with non-constant kernel solution. We will illustrate this fact by simulating differential rotation between 2 images. We keep the reference image we had already generated for the previous simulations, and make another one by rotating the frames with respect to its center with an amplitude of 0.7 pixel from one corner to the other corner of the frame. The results of subtraction are presented in Fig. 4. The systematic pattern due to rotation appears clearly in the image with a constant kernel solution, while it is completely removed in the fit of a solution with a spatial variation of order 2. This is well confirmed by the χ^2 analysis which shows that an optimal result has been reached with the non-constant kernel solution (see Fig. 5).

5.3. Under-sampling.

One last problem that can be encountered in astronomical images is under-sampling. To test the sensitivity of the method to under-sampling, we simulate a pair of very under-sampled images. We take $\alpha = 2.0$ ($FWHM = 1.17$ pixels) for the reference and $\alpha = 1.0$ ($FWHM = 1.67$ pixels) for the other image. Since our goal is just to test the effect of under-sampling only, we perform image subtraction with constant kernel solution. The resulting normalized χ^2 distribution is presented in Fig. 6. The result is as good as in the previous simulations. One may wonder why under-sampling does not induce any problems, since under-sampling should affect the convolution with the basis vectors. It is true that individually the least-square

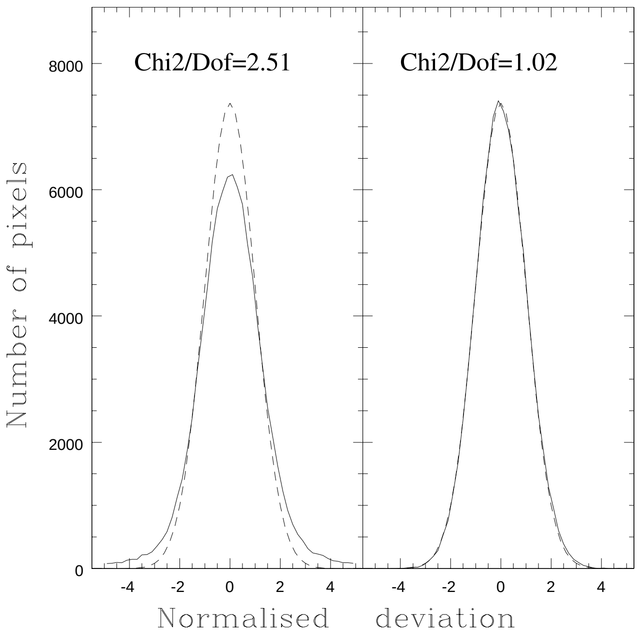

Fig. 5. Histograms of the normalized deviations in the subtracted images presented in Fig. 4. Left is the histogram for constant kernel solution, and right is the histogram for a fit of kernel variation to order 2. The dashed curve is gaussian with $\sigma = 1$. Note the goodness and the dramatic reduction of the χ^2 when fitting the kernel variations.

vectors which are obtained with under-sampled images will differ from the well sampled vectors. However one should not forget that the best solution is constructed by taking linear combinations of the least-square vectors. Even if the vectors are slightly different from the well sample vectors, an optimal linear combination of these vectors can still be made.

6. Testing with astronomical images.

It has already been shown by Alard & Lupton (1998) that nearly optimal results (χ^2/Dof = 1.05) can be achieved within small sub-areas of crowded fields. This of course breaks down as the area is enlarged, since kernel variations are no longer negligible. Comparable results should be achievable but for somewhat larger regions if a variable kernel solution is used.

We test this assumption by extracting 2 larger sub-areas (256×256) from the fields already used in Alard & Lupton (1998), and making image subtraction with constant and variable kernel solutions. The results are shown in Fig. 7 and Fig. 8. The constant kernel solution achieved only χ^2/Dof = 1.19, with numerous systematic residuals near the edges of the field. In contrast the subtracted image achieved with a variable kernel (χ^2/Dof = 1.04) is very close to the χ^2 obtained with constant kernel solution for a smaller sub-area (128×256). This analysis demonstrates the ability of the new method to deal with kernel variations in crowded-field images. It is certainly useful in this case, since larger areas and thus slightly more robust and reliable results can be obtained in crowded fields. Furthermore

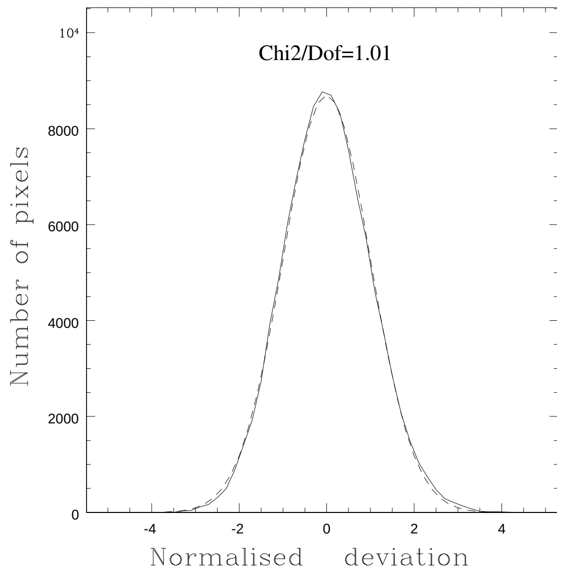

Fig. 6. Histograms of the deviations in the subtracted image in the case of very undersampled PSF's. The PSF in the reference image is very narrow with: $FWHM = 1.17$ pixels, thus we are in presence of a bad case of under-sampling. For the other image $FWHM = 1.67$, consequently it is also under-sampled.

the ability to deal with kernel variations is absolutely essential when one has to deal with fields having a low density of bright objects. This is often the case for supernovae searches, Cepheid surveys in other galaxies, and for monitoring of gravitational lenses. Thus the method has many important applications. An example can be found in the analysis of a series of images of the Huchra Lens gravitational lens (Wozniak *at al.* 2000 1998).

7. Conclusion and summary.

It has been demonstrated that image subtraction with a space-varying kernel solution could be achieved at a cost of computing time only 20 % to 30 % greater than for a constant kernel solution.

We have generated a series of Monte-Carlo images of a simulated stellar field to test the method. The ability of the method to deal with PSF variations between images has been demonstrated. It was found that even large relative PSF variations are very well corrected by the method, resulting in a χ^2/Dof very close to 1 in the subtracted image. varying kernel solutions can also automatically correct for imperfect registration between the images. This ability was illustrated by generating 2 images with differential rotation with respect to each other. While the systematic pattern due to differential rotation was visible in the subtracted image obtained with a constant kernel solution, it disappeared completely when fitted by a spatially varying kernel. To complete the Monte-Carlo analysis, we tackled the problem of under-sampling by generating 2 very under-sampled images. The χ^2 analysis showed that even

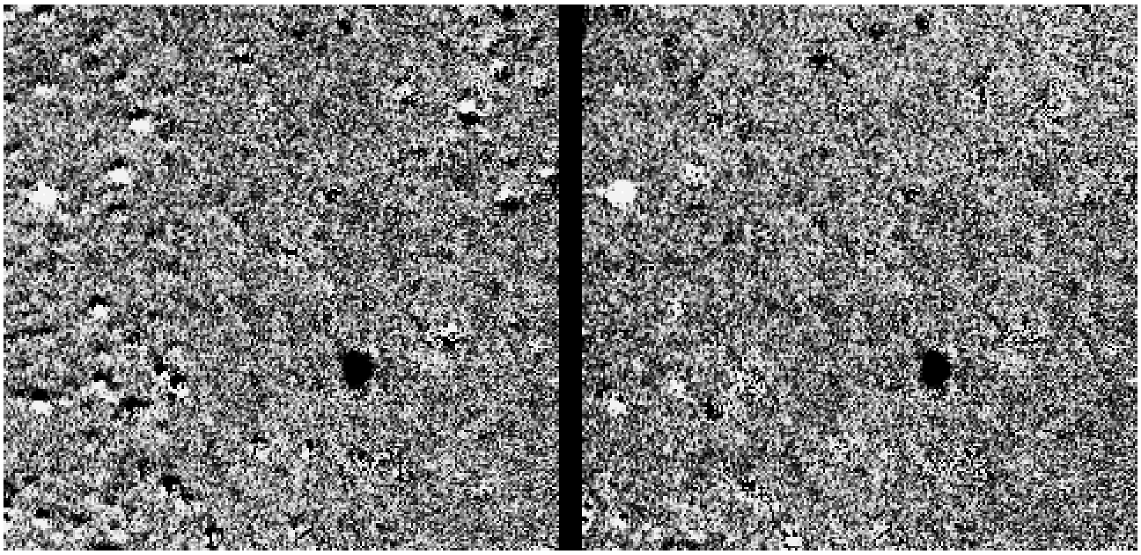

Fig. 7. Left is the subtracted image obtained with constant kernel solution. Note the systematic residuals around the bright objects on the left side of the image and in the upper right corner. Note the diseaperance of these pattern in the next image, which has been obtained by fitting the spatial variations of the kernel to order 2. Two variables are present in the field (the bright and dark spots) , and circular areas around these objects had to be excluded for χ^2 evaluation.

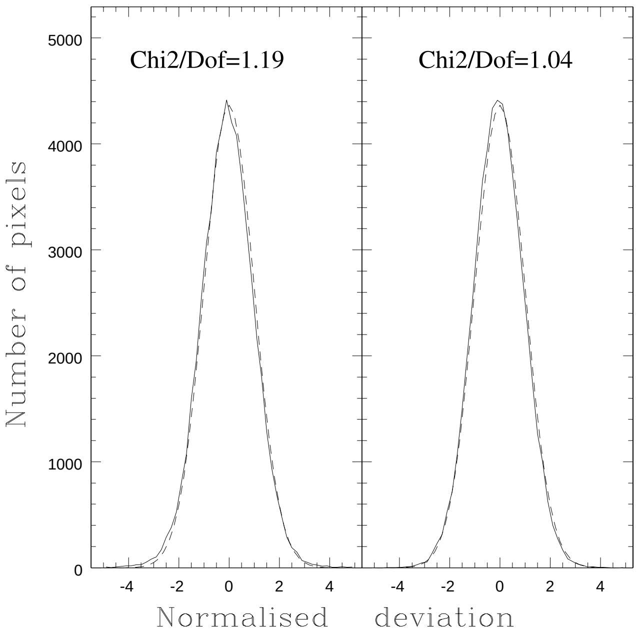

Fig. 8. Histograms of the normalized deviations in the subtracted images presented in Fig. 7. Left is the histogram for constant kernel solution, and right is the histogram for a fit of kernel variation to order 2. The dashed curve is gaussian with $\sigma = 1$.

in this case an optimal result can be achieved. Finally we applied the method to real astronomical data. Again we achieved nearly optimal results with a spatially-varying solution, even when the size of the sub-area used for the fit is increased.

These results are certainly useful for crowded field images, but have proved to be absolutely essential in case of fields with low density of high signal to noise objects. A good example is the case of the photometry of the 4 images of the Huchra lens (Wozniak *et al.* 2000 1999).

To conclude, we can say that this new extension of the image subtraction method will certainly prove to be very useful for the supernovae searches that are currently underway. It may also prove useful for variability searches in other galaxies, surveys of variables near globular cluster cores (Olech *et al.* 1999), and may also become one of the favorite methods to analyze microlensing variability in gravitational lens systems.

Software availability:
The ISIS image subtraction package is available from a web site: **http://www.iap.fr/users/alard/package.html**
This package has utilities to make image registration, image subtraction, and photometry of the images.

8. Acknowledgments

It is a pleasure to thank B. Paczyński for his support during my stay in Princeton. I thank J. Pritchard for his kind help, and especially for reading this manuscript and improving the English. I would like to acknowledge interesting discussions with B. Paczynski, P. Wozniak, and R. Lupton. I would also like to thank Dr. A. Udalski and the whole OGLE team for making their data avaiable prior to publication. It is a pleasure to thank the referee for his help with the manuscript.

References

Alard 1999, A&A, **343**, 10
Alard & Lupton 1998, ApJ, **503**, 325
Alard, C. & Guibert, J. A&A 1997, **326**, 1

Alcock, C. et al. 1993, Nature, **365**, 621
Aubourg, E. et al. 1993, Nature, **365**, 623
Olech *et al.* 1999, MNRAS, **310**, 759
Schechter, P., Mateo, M., Saha, A., M. PASP **105**, 1342 (1993)
Tomaney, A. and Crotts , A., AJ 1996 **112** 2872
Wozniak *et al.* , 2000, ApJ, **529**, 88