PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **120**:449–464, 2008 April

Optimal Image Subtraction Method: Summary Derivations, Applications, and Publicly Shared Application Using IDL

J. PATRICK MILLER,[1] C. R. PENNYPACKER,[2] AND GRAEME L. WHITE,

Centre for Astronomy, James Cook University, Townsville, Australia; pmiller@hsutx.edu

Received 2007 December 10; accepted 2008 February 20; published 2008 April 11

ABSTRACT. To detect objects that vary in brightness or spatial coordinates over time, C. Alard and R. H. Lupton in 1998 proposed an "optimal image subtraction" (OIS) method that constructs a convolution kernel from a set of matching stars distributed across the two images to be subtracted. Using multivariable least squares, the kernel is derived and can be designed to vary by pixel coordinates across the convolved image. Local effects in the optics, including aberrations or other spatially sensitive perturbations to a perfect image, can be mitigated. This paper presents the specific systems of equations that originate from the OIS method. Also included is a complete description of the Gaussian components basis vectors used by Alard & Lupton to construct the convolution kernel. An alternative set of basis vectors, called the delta function basis, is also described. Important issues are addressed, including the selection of the matching stars, differential background correction, constant photometric flux, contaminated pixel masking, and alignment at the subpixel level. Computer algorithms for the OIS method were developed, written using the Interactive Data Language (IDL), and applications demonstrating these algorithms are presented.

Online material: color figures

1. INTRODUCTION

Time-varying objects include a growing list of examples such as active galactic nuclei (AGN), variable stars, supernovae (SNe), extrasolar planets undergoing transit, microlensing events, and gamma ray bursts. Included in the list are objects that change their position in space such as asteroids, comets, Kuiper Belt objects, and SNe light echoes.

Searches for time-varying and position-changing objects are undertaken by acquiring an image of a particular region of the sky followed by a second image taken at a later time. Ideally, both images are taken with the same telescope using the same filter and CCD. The two images are aligned pixel-by-pixel then subtracted to reveal any changes in light (i.e., residual light) that occurred during the time interval between the two images.

Many subtraction methods use a convolution procedure (Phillips & Davis 1995; Tomaney & Crotts 1996; Irwin & Irwin 2003). Identifying the two images as I and R, a convolution kernel K is constructed. The R image, typically the better-seeing of the two, is convolved to match the worse-seeing image I. The convolution $R\otimes K$ is determined, and a subtraction S is given by $S = I - (R\otimes K)$. For a kernel K of size $L \times L$, where L is odd, the convolution is defined as

$$(R\otimes K)_{i,j} = \sum_{u=i-\Delta w}^{i+\Delta w} \sum_{v=j-\Delta w}^{j+\Delta w} R_{u,v} K_{u-i+\Delta w+1, v-j+\Delta w+1}, \quad (1)$$

where $\Delta w = \frac{1}{2}(L-1)$ and (i,j) represents the ij^{th} pixel of the convolution.

Alard & Lupton (1998) proposed an "optimal image subtraction" (OIS) method that constructs a convolution kernel K from a set of matching stars (they call "stamps"), found on the two images to be subtracted. Given a set of basis vectors $\{K_n:\ n = 1, 2, \ldots, N\}$ defined in § 2.3, the set of stamps is used to build a kernel $K = \sum_{n=1}^{N} a_n K_n$. The following equation is defined over the stamps:

$$I = (R\otimes K) = R\otimes\left(\sum_{n=1}^{N} a_n K_n\right) = \sum_{n=1}^{N} a_n (R\otimes K_n), \quad (2)$$

which is then solved in the least-square sense for $\{a_n:\ n = 1, 2, \ldots, N\}$, the coefficients used in the basis vector expansion for K.

$R\otimes K$ is a least-squares match to I, but not equal to I as R and I contain a time-varying or position-changing object. The subtraction is given by $S = I - (R\otimes K)$, revealing the residual light from that object because the other light sources subtract into the background.

To illustrate the procedure, two corresponding stamps of a nonvariable star are shown in Figure 1. These were taken from

[1] Department of Mathematics, Hardin-Simmons University, Abilene, TX

[2] Lawrence Berkeley National & Space Sciences Laboratory Berkeley, CA

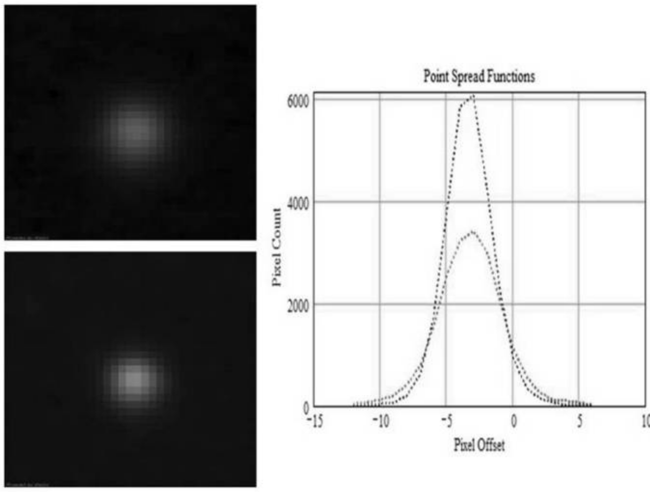

FIG. 1.—Two corresponding stamps and their point-spread functions, the first from the image and the second from the reference. Note that the point-spread functions are unmatched. See the electronic edition of the PASP for a color version of this figure.

an image and reference, in which there was a galaxy with a suspected AGN. Also shown in Figure 1 are the point-spread functions (PSFs) of the two stamps. These functions are clearly unmatched.

However, Figure 2 shows the image stamp compared to the convolved reference stamp, derived by using the OIS method. In this case, the point-spread functions are matched. Figure 3 shows the subtraction $S = I - (R \otimes K)$ and the difference of the two PSFs. The average difference is 13.1 compared to the mean background sky of -1.44 with standard deviation

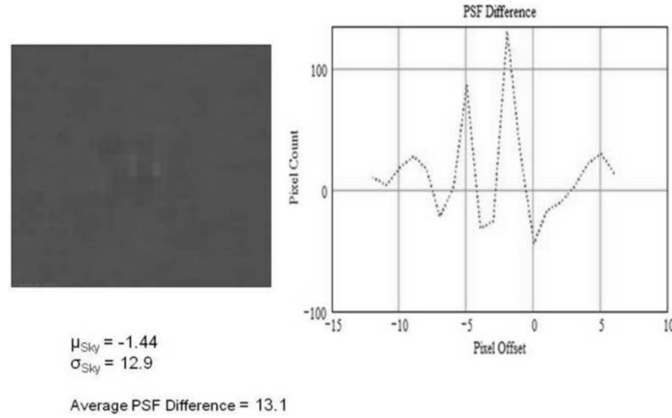

FIG. 3.—Subtraction of the two stamps in Figure 2. Also shown is the difference between the point-spread functions of the same two stamps. The star in the stamps is subtracted into the background. See the electronic edition of the PASP for a color version of this figure.

of 12.9. That is, the average PSF of the subtracted stamp is at 1 σ of the mean background sky, meaning the stamp of the nonvariable star has subtracted into the background.

A stamp of the galaxy with a suspected AGN is shown in Figure 4. This stamp was taken from the same image and reference as the stamp of the nonvariable star in Figure 1. Note that the point-spread function of the galaxy from the image does not match that from the convolved reference as it did in the case of the star.

As shown in Figure 5, the average of the PSF of the subtracted stamp is at 67 σ of the mean background sky. Therefore, the subtraction of the two does not go into the background but

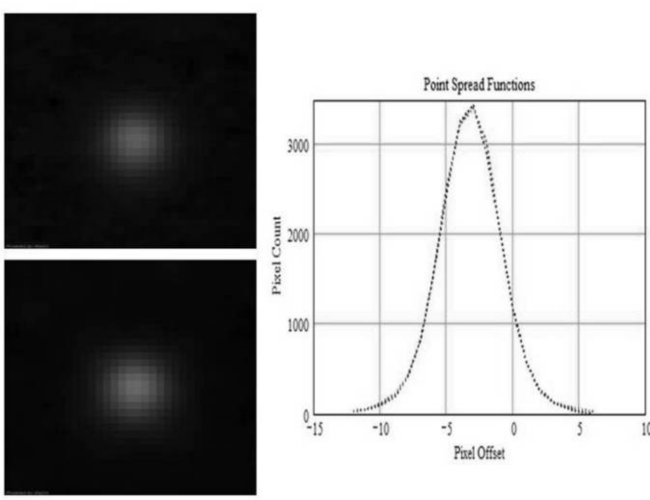

FIG. 2.—Two corresponding stamps and their point-spread functions, the first from the image and the second from the convolved reference using the OIS method. In this instance, note that the point-spread functions are matched. See the electronic edition of the PASP for a color version of this figure.

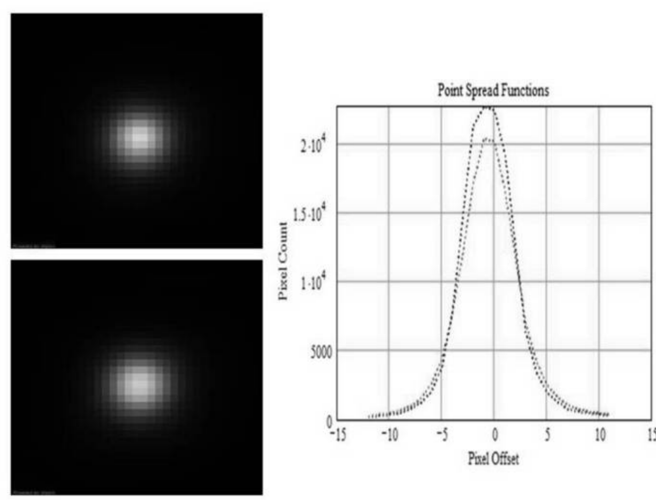

FIG. 4.—Two stamps of a galaxy with a suspected active nucleus, the first from the image and the second from the convolved reference. Note the point-spread functions are not matched. See the electronic edition of the PASP for a color version of this figure.

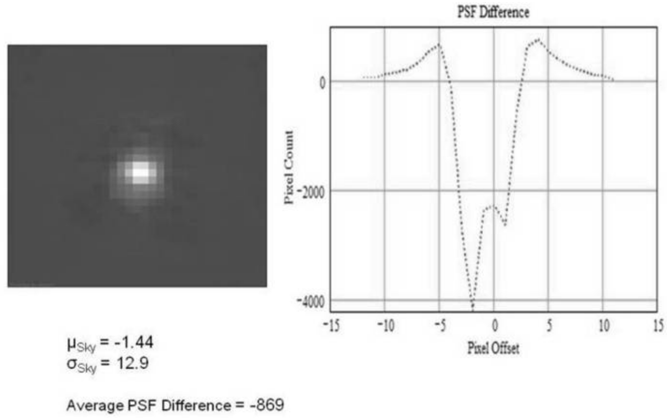

FIG. 5.—Subtraction of the stamps of the galaxy in Figure 4, one convolved to match the other in the least-squares sense using the OIS method. The galaxy in the stamps is not subtracted into the background. There is a change in the galaxy from the image to the reference, during the time interval between the two. See the electronic edition of the PASP for a color version of this figure.

reveals changes in the galaxy during the time interval (one year) between the image and the reference, verifying the status of the galaxy as an AGN candidate.

A number of authors have provided a broad description or overview of the mathematics required by the OIS method (Alard 2000; Bond, et al. 2001; Astier 2004, private communication; Israel et al. 2007). The literature does not contain the specific systems of equations that originate from the least-squares method or a complete description of the basis vectors used to construct the convolution kernel K. Both of these are shown in subsequent sections of this paper.

In addition, computer algorithms for the OIS method were developed from these systems of equations, and written using the Interactive Data Language (IDL) from ITT Visual Information Solutions, Inc. (Boulder, CO), a Fortran-like interpretive language used in data analysis and image processing. These algorithms are now available for increasing public distribution and use. Successful applications of these IDL computer algorithms are presented in subsequent sections, validating the algorithms and demonstrating the flexibility of the OIS method to detect the various kinds of time-varying and position-changing objects.

2. OPTIMAL IMAGE SUBTRACTION METHOD (OIS)

The following shows the overall flow of the OIS method. A subset of the image I represented as a collection of stamps across the entire image is used to derive the convolution kernel K:

1. Let I and R be the aligned image and reference. R is the better-seeing image. The two images are of equal dimension $S_R x S_C$.

2. Select the stamps containing the brighter, unsaturated stars $\{I_k:\ k = 1, 2, \ldots, P\}$ within the image I. $\{R_k:\ k = 1, 2, \ldots, P\}$ are the corresponding stamps within the reference R.

3. Select the basis vectors for the $L \times L$ kernel K, $\{K_n:\ n = 1, 2, \ldots, N\}$ where $1 \le N \le L^2$.

4. Using the stamps and the multivariable least-squares method, solve the equation $I = R \otimes K = \sum_{n=1}^{N} a_n (R \otimes K_n)$ for $\{a_n:\ n = 1, 2, \ldots, N\}$.

5. Conduct a χ^2 fitness test on each stamp $\{I_k - (R_k \otimes K):\ k = 1, 2, \ldots, P\}$, where $K = \sum_{n=1}^{N} a_n K_n$.

6. Reject the failing stamps.

7. If no stamps are rejected, proceed to the next step. Otherwise, go to the multivariable least-squares method and repeat the calculation with the reduced set of stamps.

8. Subtract the images $S = I - (R \otimes K)$.

2.1. Constant Kernel

Consider the equation $I = R \otimes K$ to be solved in the least-squares sense over the stamps. Let $E_{k,i,j}$ be defined as the following where $\eta_k \times \eta_k$ is the size of the k^{th} stamp, assumed to be odd:

$$E_{k,i,j} = (R_k \otimes K)_{i,j} - (I_k)_{i,j} = \left(R_k \otimes \sum_{n=1}^{N} a_n K_n \right)_{i,j} - (I_k)_{i,j}$$
$$= \sum_{n=1}^{N} a_n (R_k \otimes K_n)_{i,j} - (I_k)_{i,j}, \tag{3}$$

where $1 \le k \le P$, $1 \le i \le \eta_k$, and $1 \le j \le \eta_k$. The least-squares function is given by

$$F(a_1, a_2, \ldots, a_N) = \sum_{k=1}^{P} \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k} (E_{k,i,j})^2 = \sum_{k=1}^{P} \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k}$$
$$\times \left[\sum_{n=1}^{N} a_n (R_k \otimes K_n)_{i,j} - (I_k)_{i,j} \right]^2. \tag{4}$$

The least-squares solution is$\nabla F = 0$. This produces the $N \times N$ linear system

$$Ma = b, \tag{5}$$

where $M_{n,l} = \sum_{k=1}^{P} \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k} (R_k \otimes K_n)_{i,j} (R_k \otimes K_l)_{i,j}$ and $b_l = \sum_{k=1}^{P} \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k} (I_k)_{i,j} (R_k \otimes K_l)_{i,j} n = 1, 2, \ldots, N$ and $l = 1, 2, \ldots, N$.

The solutions $\{a_n : n = 1, 2, \ldots, N\}$ consist of the coefficients that define the best-fit kernel $K = \sum_{n=1}^{N} a_n K_n$. In this instance, the kernel remains constant over the image (i.e., does not vary as a function of pixel coordinates).

2.2. Space-varying Kernel

To allow for the possibility that the kernel K varies across the image, the coefficients a_n of the expansion are redefined as a_n

(x, y) where (x, y) are pixel coordinates. As proposed by Alard (2000) it is assumed that these coefficients are bivariate polynomials of (x, y); each $a_n(x, y)$ multiplies a basis vector K_n:

$$a_n(x,y) = \sum_{m=0}^{d}\sum_{l=0}^{d-m} \alpha_{n,m,l} x^m y^l, \tag{6}$$

$$K(x,y) = \sum_{n=1}^{N} a_n(x,y) K_n, \tag{7}$$

where d is the degree of the polynomial. $a_n(x, y)$ contains $\frac{1}{2}(d+1)(d+2)$ terms. The degree of the polynomial defines the degree of the OIS (e.g., a first degree OIS uses a bivariate polynomial with $d = 1$).

As proposed by Alard (2000), on each stamp $a_n(x, y)$ is assumed to be constant and set to its value at the coordinates of the center pixel $(\hat{x}_k, \hat{y}_k)$ of the stamp:

$$\begin{aligned} E_{k,i,j} &= \sum_{n=1}^{N} a_n (R_k \otimes K_n)_{i,j} - (I_k)_{i,j} \\ &= \sum_{n=1}^{N}\left[\sum_{m=0}^{d}\sum_{l=0}^{d-m} \alpha_{n,m,l} \hat{x}_k^m \hat{y}_k^l (R_k \otimes K_n)_{i,j}\right] - (I_k)_{i,j}. \end{aligned} \tag{8}$$

The least-squares function is given by

$$\begin{aligned} F = \sum_{k=1}^{P}\sum_{i=1}^{\eta_k}\sum_{j=1}^{\eta_k} (E_{k,i,j})^2 = \sum_{k=1}^{P}\sum_{i=1}^{\eta_k}\sum_{j=1}^{\eta_k} \\ \times \left\{\sum_{n=1}^{N}\left[\sum_{m=0}^{d}\sum_{l=0}^{d-m} \alpha_{n,m,l} \hat{x}_k^m \hat{y}_k^l (R_k \otimes K_n)_{i,j}\right] - (I_k)_{i,j}\right\}^2, \end{aligned} \tag{9}$$

and the least-squares solution is given by$\nabla F = 0$.

This produces a $Q \times Q$ linear system where $Q = \frac{1}{2}N(d+1)(d+2)$:

$$\begin{aligned} \sum_{n=1}^{N}\sum_{m=0}^{d}\sum_{l=0}^{d-m} \alpha_{n,m,l} \Bigg[\sum_{k=1}^{P}\sum_{i=1}^{\eta_k}\sum_{j=1}^{\eta_k} \hat{x}_k^{m+r} \hat{y}_k^{l+s} (R_k \otimes K_n)_{i,j} \\ \times (R_k \otimes K_q)_{i,j}\Bigg] = \sum_{k=1}^{P}\sum_{i=1}^{\eta_k}\sum_{j=1}^{\eta_k} \hat{x}_k^r \hat{y}_k^s (I_k)_{i,j} (R_k \otimes K_q)_{i,j}, \end{aligned} \tag{10}$$

$q = 1, 2, \ldots, N$ and for each q, $r = 0, 1, \ldots, d$, $s = 0, 1, \ldots, d - r$.

This is a more difficult linear system to visualize than the system for the constant kernel. A 6×6 example for $N = 2$ and $d = 1$ is provided in Appendix A.

The solutions $\{\alpha_{n,m,l} : n = 1, 2 \ldots, N m = 0, 1, \ldots, d l = 0, 1, \ldots, d - m\}$ consist of coefficients that define the best-fit kernel $K(x, y) = \sum_{n=1}^{N}(\sum_{m=0}^{d}\sum_{l=0}^{d-m} \alpha_{n,m,l} x^m y^l) K_n$. This kernel varies as a function of the image pixel coordinates (x, y).

2.3. Basis Vectors

The basis vectors for the $L \times L$ kernel K are defined as $\{K_n : n = 1, 2, \ldots, N\}$, where L is odd and $1 \le N \le L^2$. Two different bases were considered, the delta function basis (DFB) and Gaussian components basis (GCB).

The DFB was defined for the full L^2 dimension of the kernel. A basis vector is an $L \times L$ matrix with 0's in all entries except in one entry there is 1. The 1 varies across all of the L^2 entries to produce the basis.

The formal definition of the DFB is $(K_n)_{i,j} = \delta_{n,L(i-1)+j}$, where $n = 1, 2, \ldots, L^2$ while $i = 1, 2, \ldots, L$, and $j = 1, 2, \ldots, L$ for each value of n.

$$\delta_{a,b} = \begin{cases} 0 & \text{if } a \ne b \\ 1 & \text{if } a = b \end{cases}$$

is the Kronecker delta. Because L is odd, the centered delta function basis vector is included, allowing for two images with almost-identical stamps.

The GCB is not defined for the full L^2 dimension of the kernel. Instead, it is defined over an N-dimensional subspace as follows (Alard 2000; Astier 2004, private communication):

1. N_c is the number of Gaussian components, and defines the number of degrees $\{d_k : k = 1, 2, \ldots, N_c\}$ and the number of standard deviations $\{\sigma_{\text{Kernel}}\sigma_k : k = 1, 2, \ldots, N_c\}$ of the components. σ_{Kernel} is the standard deviation of the kernel K:

2. Each component contributes a total of $\frac{1}{2}(d_k + 1)(d_k + 2)$ matrices $(L \times L)$ to the basis vectors. The total number of basis vectors is

$$N = 1 + \sum_{k=1}^{N_c} \frac{1}{2}(d_k + 1)(d_k + 2). \tag{11}$$

3. Let $Ctr = \frac{1}{2}(L + 1)$ then for $n = 1, 2, \ldots, N - 1$, where each combination of u and v produces one basis vector while $0 \le u \le d_k$ and $0 \le v \le d_k - u$:

$$(K_n)_{i,j} = (i - Ctr)^u (j - Ctr)^v e^{-[(i-Ctr)^2 + (j-Ctr)^2]/2\sigma_{\text{Kernel}}^2 \sigma_k^2}. \tag{12}$$

$1 \le i \le L$ and $1 \le j \le L$ and

$$(K_N)_{i,j} = \begin{cases} 0 & \text{if } i \ne Ctr \vee j \ne Ctr \\ 1 & \text{if } i = Ctr \wedge j = Ctr \end{cases}. \tag{13}$$

The purpose of including K_N is to allow for images that are almost identical. This is the centered delta function basis vector that just returns the pixel value (i.e., no weighted average) during the convolution R (i.e., $R \otimes K_N = R$).

As per Astier (2004, private communication) the following Gaussian components were selected:

$$N_c = 3, \qquad d = (6, 4, 2), \qquad \sigma = (0.7, 1.5, 2.0). \tag{14}$$

Astier established these components empirically, and these values are similar to those proposed by Alard (2000). More recently in an effort to eliminate empirical estimates, Israel et al. (2007) developed a procedure for selecting optimal components for image sets with different S/N ratios and sampling.

2.4. Differential Background

Alard & Lupton (1998) proposed that the differential background between I and R can be modeled with a bivariate polynomial $p(x,y) = \sum_{k=0}^{\hat{d}} \sum_{l=0}^{\hat{d}-k} c_{k,l} x^k y^l$ added to the equation $R \otimes K(x,y) = I + p(x,y)$ where $\hat{d}$ is the degree of the polynomial and (x,y) are the pixel coordinates. Solving this equation in the least-squares sense produces a linear system that includes the coefficients both of the basis vectors for the space-varying kernel K and of the differential background polynomial $p(x,y)$.

In the IDL computer algorithms, the differential background was processed separately from the best-fit kernel calculation. That is, the differential background was corrected first on I and R then the best-fit kernel was derived using the corrected version. A bivariate polynomial $p(x,y)$ was used to model the background, but derived independently of the kernel calculation using a least-squares procedure over $B = \{I_{i,j} : |I_{i,j} - \mu| \le \sigma\}$, the set of pixels in the image I within 1 σ of the mean sky. For a derivation of the polynomial $p(x,y)$ refer to Appendix B.

2.5. Constant Photometric Ratio

It is essential for many astronomical applications to conserve the photometric flux when using the convolution kernel. That is, the photometric flux of the convolved reference $R \otimes K$ is equal to that of the image I being matched. This is not a difficult thing to do once it is realized that the integral of the kernel (i.e., the sum of the entries in the kernel) $\|K\| = \sum_{l=1}^{L} \sum_{m=1}^{L} K_{l,m}$ is the scaling factor of the photometric flux $\Im$ from the reference stamp R_k to the convolved reference stamp $R_k \otimes K$:

$$\Im_{R_k \otimes K} = \|K\| \Im_{R_k}, \tag{15}$$

where $1 \le k \le P$, $\Im_{R_K} = \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k} (R_k)_{i,j}$, and $\Im_{R_K \otimes K} = \sum_{i=1}^{\eta_k} \sum_{j=1}^{\eta_k} (R_k \otimes K)_{i,j}$. In other words, in order for the flux of the reference stamp R_k to equal the flux of the convolved reference stamp $R_k \otimes K$, it is necessary that the kernel integral $\|K\| = 1$.

This can be done by redefining the kernel basis vectors $\{K_n : n = 1, 2, \ldots, N\}$ in the following manner (Alard 2000). First, for $n = 1, 2, \ldots, N$ normalize the basis vectors:

$$\tilde{K}_n = \begin{cases} \frac{1}{\|K_n\|} K_n & \text{if } \|K_n\| \neq 0 \\ K_n & \text{if } \|K_n\| = 0 \end{cases} \|\tilde{K}_n\|$$

$$= \sum_{l=1}^{L} \sum_{m=1}^{L} (\tilde{K}_n)_{l,m} \begin{cases} 1 & \|K_n\| \neq 0 \\ 0 & \|K_n\| = 0 \end{cases}. \tag{16}$$

Second, pick any one of the basis vectors $\tilde{K}_u$ such that $\|\tilde{K}_u\| = 1$. A convenient choice for this vector is the centered delta function basis vector. Third, define a new basis $\{\hat{K}_n | n = 1, 2, \ldots, N\}$ as follows:

$$\hat{K}_1 = \tilde{K}_u \quad \text{for } n = 1, \tag{17}$$

and for $n = 2, 3, \ldots, N$

$$\hat{K}_n = \begin{cases} \tilde{K}_n - \hat{K}_1 & \text{if } \|\tilde{K}_n\| = 1 \\ \tilde{K}_n & \text{if } \|\tilde{K}_n\| = 0 \end{cases}, \tag{18}$$

$$\|\hat{K}_n\| = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}. \tag{19}$$

If the best-fit space-varying kernel $K(x,y) = \hat{K}_1 + \sum_{n=2}^{N} a_n(x,y) \hat{K}_n$ is solved on this basis, its integral equals 1 over the entire image. For a proof of this statement, refer to Appendix C

Consider the following photometric ratio of stamps I_k to R_k and constraint upon the best-fit kernel:

$$\xi \equiv \frac{\Im_{I_k}}{\Im_{R_k}} \qquad \hat{K}(x,y) = \xi \hat{K}_1 + \sum_{n=2}^{N} a_n(x,y) \hat{K}_n. \tag{20}$$

In this case, finding the best-fit space-varying kernel will produce a convolved reference with the same flux as the image, preserving the photometric ratio of the kth stamp.

In the computer algorithms written for this study $\xi = median\left\{\frac{\Im_{I_k}}{\Im_{R_k}} : \ k = 1, 2, \ldots, P\right\}$. This gives an best-fit space-varying kernel that approximately conserves the photometric ratios of all of the stamps.

3. ADDITIONAL ISSUES

3.1. Identifying and Masking Bad Pixels

In addition to the mathematics of the OIS, a number of additional issues need to be addressed that are important to a successful application of the method. One such issue is the selection of stamps (i.e., the set of matching stars from which the best-fit kernel is derived).

Images contain bad pixels such as hot pixels, cosmic ray hits, and saturated pixels including extensive vertical or horizontal saturation. In the selection procedure for the stamps, these bad pixels must be excluded from consideration. Of course the stamps can be hand selected; however, it is possible to mask the bad pixels and automate the selection of a set of uniformly distributed stamps.

One algorithm to process bad pixels is given by the following:

1. Identify bad pixels in I and R, separately.
2. Flag pixels from I and R that are bad in either I or R.
3. Determine the background sky, mean μ and standard deviation σ for both I and R.
4. Replace the flagged pixels with random values of background sky within 1 σ of μ in both I and R.

The resulting masking of I and R forms "ghost" images, I_S and R_S. It is from I_S and R_S that the stamps are selected, and not from I and R.

An automated procedure to construct I_S and R_S was written using IDL that incorporated the *Sky* utility from the IDL Astronomy User's Library maintained by the NASA Goddard Space Flight Center (GSFC; Greenbelt, MD). *Sky* calculates the background statistics of an image including the mean and standard deviation. See Landsman (2004).

The background sky mean μ and standard deviation σ are determined for both I and R. The following images are constructed:

$$I_Q = \frac{|I - \mu_I|}{\sigma_I} \qquad R_Q = \frac{|R - \mu_R|}{\sigma_R}. \tag{21}$$

At each pixel in I_Q and R_Q the number of standard deviations from the mean sky is recorded. The minimum of the median and standard deviation of the pixels in I_Q and R_Q is calculated:

$$\begin{aligned} \mathrm{M}_Q &= \min[\mathrm{median}(IQ), \mathrm{median}(RQ)], \\ \Sigma_Q &= \min[\sigma_{IQ}, \sigma_{RQ}]. \end{aligned} \tag{22}$$

A range of light $R_{\min}$ to $R_{\max}$ is set in which the stamps are selected (i.e., the pixels not within this range are masked to form I_S and R_S):

$$\begin{aligned} F_{\min} &= R_{\min}\left(\frac{\max[\max(IQ), \max(RQ)] - \mathrm{M}_Q}{\Sigma_Q}\right), \\ F_{\max} &= R_{\max}\left(\frac{\max[\max(IQ), \max(RQ)] - \mathrm{M}_Q}{\Sigma_Q}\right), \\ g_{\min} &= \mathrm{M}_Q - |F_{\min}|\Sigma_Q, \\ f_{\min} &= \max(3, g_{\min}), \qquad f_{\max} = \mathrm{M}_Q - |F_{\max}|\Sigma_Q. \end{aligned} \tag{23}$$

Finally, the pixels that satisfy the following condition are used in the "ghost" images to find the stamps while the other pixels are masked, including any unused pixels:

$$\begin{aligned} (f_{\min}\sigma_I \le |I - \mathrm{M}_Q| \le f_{\max}\sigma_I) \wedge (f_{\min}\sigma_R \\ \le |R - \mathrm{M}_Q| \le f_{\max}\sigma_R). \end{aligned} \tag{24}$$

The following shows the empirical results in setting the parameters $R_{\min}$ and $R_{\min}$:

1. For well-behaved images with no significant amount of saturation or blooming (e.g., no linear streaks), $[R_{\min}, R_{\max}] = [0.25, 0.75]$ seems to work well. Essentially, most of the light is available for the search of stamps. To obtain all of the light $[R_{\min}, R_{\max}] = [0.00, 1.00]$

2. For significant saturation or blooming, especially with linear streaks, or extensive light from the Moon or nebulae, $[R_{\min}, R_{\max}] = [0.00, \quad 0.15]$ seems to work. Essentially, the bright, diffuse light is not available for stamps, and the stamps must be derived from the fainter stars.

3.2. Kernel Size & Selection of the Stamps

The selection of the stamps is very important. It is best that they include single stars with minimal light contamination and be uniformly distributed over I and R. Both of the IDL utilities *Sky* and *Find* from GSFC were used in a stamp selection algorithm.

From *Sky* the mean and standard deviation of the background of I_S are calculated. The standard deviation is an input parameter into *Find* that identifies the centroids of point sources on I_S.

Once identified, the centroids are individually examined. A $(2\sigma + 1)x(2\sigma + 1)$ stamp is placed around the centroid where σ is derived from the approximate FWHM of the worse-seeing image, a 2D Gaussian fit is performed, and the stamp is increased by 2×2 with a second 2D Gaussian fit performed. The size of the stamp is increased by 2×2 until the percent difference in $(\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}$ is less than a preset parameter (typically set at 0.1%).

The estimated kernel σ is calculated, σ_{Kernel} by comparing the first two stamps from I_S and R_S (i.e., the first two stamps not overlapping the edge of the image). A 2D Gaussian fit is performed on the two stamps:

$$\sigma_{\mathrm{Kernel}} = \tfrac{1}{2}(|\sigma_{I_S,1} - \sigma_{R_S,1}| + |\sigma_{I_S,2} - \sigma_{R_{S,},2}|). \tag{25}$$

The size of the $L \times L$ kernel is given by $L = 8\sigma_{\mathrm{Kernel}} + 5$ (Astier 2004, private communication), but not less than 7 or greater than 13.

The stamps are resized to accommodate the size of the kernel and then checked to see if they overlap the boundary of the image, meet the user-defined sharpness and roundness statistical criteria of *Find*, have a flux exceeding the user-defined percent of the maximum flux of the stamps, and overlap one another. Stamps that fail any one of these tests are rejected.

3.3. Subpixel Alignment

3.3.1. Subpixel Translations

Alignment of the image and reference down to the subpixel level is needed for good subtractions, especially in the search for SNe and AGN. Once the initial alignment is done, a subpixel adjustment in translation can be accomplished using spline methods or convolution methods utilizing small polynomial interpolation kernels. For the IDL computer algorithms second or fourth degree polynomial interpolation kernels were used.

To achieve a subpixel horizontal shift to the right on an image I, look at pixel $I_{i,j}$. On either side of this pixel, to the left and to the right, are the pixels $I_{i-1,j}$ and $I_{i+1,j}$. A second degree interpolating polynomial is to be passed through the pixels. To build a second degree interpolating polynomial passing through these three pixels, construct the following divided difference table:

$$\begin{array}{ccccc} i-1 & I_{i-1,j} & & & \\ & & I_{i,j}-I_{i-1,j} & & \\ i & I_{i,j} & & \frac{1}{2}(I_{i+1,j}-2I_{i,j}+I_{i-1,j}). & \\ & & I_{i+1,j}-I_{i,j} & & \\ i+1 & I_{i+1,j} & & & \end{array} \quad (26)$$

This produces the following second degree polynomial representation for the pixel value as a function of pixel position:

$$\begin{aligned} I_{u,j} - I_{i-1,j} &+ (I_{i,j} - I_{i-1,j})(u-i+1) \\ &+ \tfrac{1}{2}(I_{i+1,j} - 2I_{ij} + I_{i-1,j})(u-i+1)(u-i). \end{aligned} \quad (27)$$

This produces the following 3×3 kernel:

$$K_H = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2}(u-i) & -(u-i+1) & \frac{1}{2}(u-i+1) \\ \times(u-i-1) & \times(u-i-1) & \times(u-i) \\ 0 & 0 & 0 \end{bmatrix}. \quad (28)$$

A fourth degree interpolating polynomial using $I_{i-2,j}$, $I_{i-1,j}$, $I_{i,j}I_{i+1,j}$, and $I_{i+2,j}$ is given by the following, from which a 5×5 kernel can be obtained:

$$\begin{aligned} I_{u,j} = I_{i-2,j} &+ (I_{i-1,j} - I_{i-2,j})(u-i+2) \\ &+ \frac{1}{2}(I_{i,j} - 2I_{i-1,j} + I_{i-2,j})(u-i+2)(u-i+1) \\ &+ \frac{1}{6}(I_{i+1,j} - 3I_{ij} + 3I_{i-1,j} - I_{i-2,j})(u-i+2) \\ \times (u-i+1)(u-i) &+ \frac{1}{24}(I_{i+2,j} - 4I_{i+1,j} + 6I_{i,j} - 4I_{i-2,j}) \\ &\times (u-i+2)(u-i+1)(u-i)(u-i-1). \end{aligned} \quad (29)$$

In general, evaluating $I_{u,j}$ at $u = i + x$ where $-1 \le x \le 1$ produces subpixel horizontal shifts, to the right when $-1 \le x < 0$ and to the left when $0 < x \le 1$. Of course, there is no shift when $x = 0$.

The following two $I_{i,u}$ produce subpixel vertical shifts, 3×3 and 5×5 respectively.

A second degree polynomial, 3×3 kernel is given by

$$\begin{aligned} I_{i,u} = I_{i,j-1} &+ (I_{i,j} - I_{i,j-1})(u-j+1) \\ &+ \tfrac{1}{2}(I_{i,j+1} - 2I_{i,j} + I_{i,j-1})(u-j+1)(u-j). \end{aligned} \quad (30)$$

A fourth degree polynomial, 5×5 kernel is given by

$$\begin{aligned} I_{i,u} = I_{i,j-2} &+ (I_{i,j-1} - I_{i,j-2})(u-j+2) \\ &+ \tfrac{1}{2}(I_{i,j} - 2I_{i,j-1} + I_{i,j-2})(u-j+2)(u-j+1) \\ &+ \tfrac{1}{6}(I_{i,j+1} - 3I_{i,j} + 3I_{i,j-1} - I_{i,j-2})(u-j+2) \\ &\times (u-j+1)(u-j) + \tfrac{1}{24}(I_{i,j+2} - 4I_{i,j+1} + 6I_{i,j} \\ &- 4I_{i,j-1} + I_{i,j-2})(u-j+2)(u-j+1) \\ &\times (u-j)(u-j-1). \end{aligned} \quad (31)$$

In general, evaluating $I_{i,u}$ at $u = j + y$ where $-1 \le y \le 1$, produces subpixel vertical shifts, up when $-1 \le y < 0$, down when $0 < y \le 1$, and no shift when $y = 0$.

If x and y are chosen, K_H and K_V can be built that shifts the image horizontally and vertically, respectively. $(I \otimes K_H) \otimes K_V = (I \otimes K_V) \otimes K_H$ produces a subpixel shift of magnitude $\sqrt{x^2 + y^2}$ at a polar angle $\theta = \tan^{-1}\frac{y}{x}$, measured in the pixel coordinate system.

A proof of the commutivity of the convolution $(I \otimes K_H) \otimes K_V = (I \otimes K_V) \otimes K_H$ is found in Appendix D. Furthermore, $\|K_H\| = \|K_V\| = 1$, and this is proved for the second degree interpolating polynomial but is true in general (i.e., vertical and horizontal interpolating kernels conserve the photometric flux).

It is possible to generalize the derivation of $I_{i,u}$ for an arbitrary, odd-sized $L \times L$ kernel, which requires an interpolating polynomial of degree $L - 1$. Figures 6, 7, and 8 provide an al-

ALGORITHM TO CALCULATE THE SUB-PIXEL SHIFT KERNELS

```
Program Sub-Pixel Shift
  Input I, x, y, L
    Call Divisors (L)
    Dimension K_LxL
    m = (L-1)/2
    Call Kernel (x, L, δ)          }
    For i = 0,1,..., L-1           }
        K(i, m) = A(i)             } Horizontal Shift by |x|
    End For                        }
    I = I ⊗ K                      }
    Call Kernel (x, L, δ)          }
    Set K = 0                      }
    For j = 0,1,..., L-1           }
        K(m, j) = A(j)             } Vertical Shift by |y|
    End For                        }
    I = I ⊗ K                      }
  Output I
End Program
```

FIG. 6.—Algorithm to calculate kernels used in the subpixel translations.

```
Sub Program Divisors , L
    Dimension δ_LxL
    For i = 0,1,..., L − 1
        δ_i,i = i!
        Sign = −1
        For j = i + 1, i + 2,..., L − 1
            δ_i,j = Sign (j − i)! δ_i,i
            Sign = −Sign
        End For
    End For
    Return δ
End Sub Program
```

FIG. 7.—Subprogram to calculate divisors for the $L \times L$ subpixel translation kernel.

```
Sub Program Kernel , Shift , L, δ
    Dimension Z_L, A_L
    m = (L − 1)/2
    Z_0 = 1
    For k = 1,2,..., L − 1
        Z_k = (Shift + m) Z_k−1
        m = m − 1
    End For
    For i = 0,1,..., L − 1
        For j = i, i + 1,..., L − 1
            A_i = A_i + Z_j / δ_i,j
        End For
    End For
    Return A
End Sub Program
```

FIG. 8.—Subprogram to calculate the $L \times L$ polynomial interpolation kernel.

gorithm that will produce this arbitrary kernel. It is assumed that when the arrays are initially dimensioned the values are set to 0, and it is assumed that the array index starts at 0 and not 1, which is consistent with IDL.

3.3.2. Corrections with the OIS Method

The OIS method is forgiving when there are systematic errors in alignment at the subpixel level. If the error is a translation, the problem can be resolved using a first-degree or even a constant-kernel OIS (Miller 2007). If there are differential rotation errors, a second-degree OIS can resolve the problem (Alard 2000).

Figure 9 shows the subtractions of a simulated image and reference, both $1\text{K} \times 1\text{K}$ in size. A total of 100 stars were randomly distributed, with FWHM of 5 for the image stars and 4 for the corresponding reference stars. The peak flux was uniformly distributed from 0 to 2×10^5 with a random Poisson background added, deviated from 60 on the image and 20 on the reference.

In Figure 9 the OIS on the left was done with the Gaussian components basis proposed by Astier (2004, private communication), and using a constant kernel. On the right is an OIS with the delta function basis, also using a constant kernel. The systematic translation errors of $\Delta x = 1$ and $\Delta y = -2$ are clearly seen in the OIS on the left while the OIS on the right shows that even a constant kernel can correct these errors.

Again using a simulation with 100 randomly distributed stars, the reference was rotated counterclockwise by 0.05° with respect to the image. Figure 10 shows a subtraction done with the OIS method, using the DFB and constant kernel. The rotation is easily seen. Figure 11 shows on the left a first degree OIS and a second degree on the right. The first degree does not quite correct the rotation effect but it is corrected in the second degree.

3.4. Contaminant Masking

In the optimal image subtractions, the bad pixels remain as contaminants, often making it difficult to identify the object being sought. One example of this occurs when looking for moving objects within crowded star fields. The mis-subtractions of the saturated stars can easily populate the subtraction. The following is a procedure that masks these bad pixels, so that the moving objects are easier to detect:

1. Use the image and reference, I and R
2. Subtract $S = I - R \otimes K$
3. Identify pixels on I that are 3 σ above the mean sky of I
4. Identify pixels on R that are 3 σ above the mean sky of R
5. Identify pixels on S that are 3 σ above the mean sky of S
6. Identify the above pixels that are common to I, R, and S
7. Mask the identified common pixels on S with the mean sky of S

In other words, bright objects on I and R that remain bright on S will be masked. Objects that move will not appear bright on I and R at the same locations, and thus not be masked but instead highlighted on the subtraction. Even though hot pixels and cosmic ray hits remain as they are random across the completed subtraction. Typically, they do not exhibit the same signature as that of a moving object.

A heavy-handed masking can be performed by only finding the pixels common to I and R instead of I, R, and S. This masking is used only in rare, and truly undesirable, situations

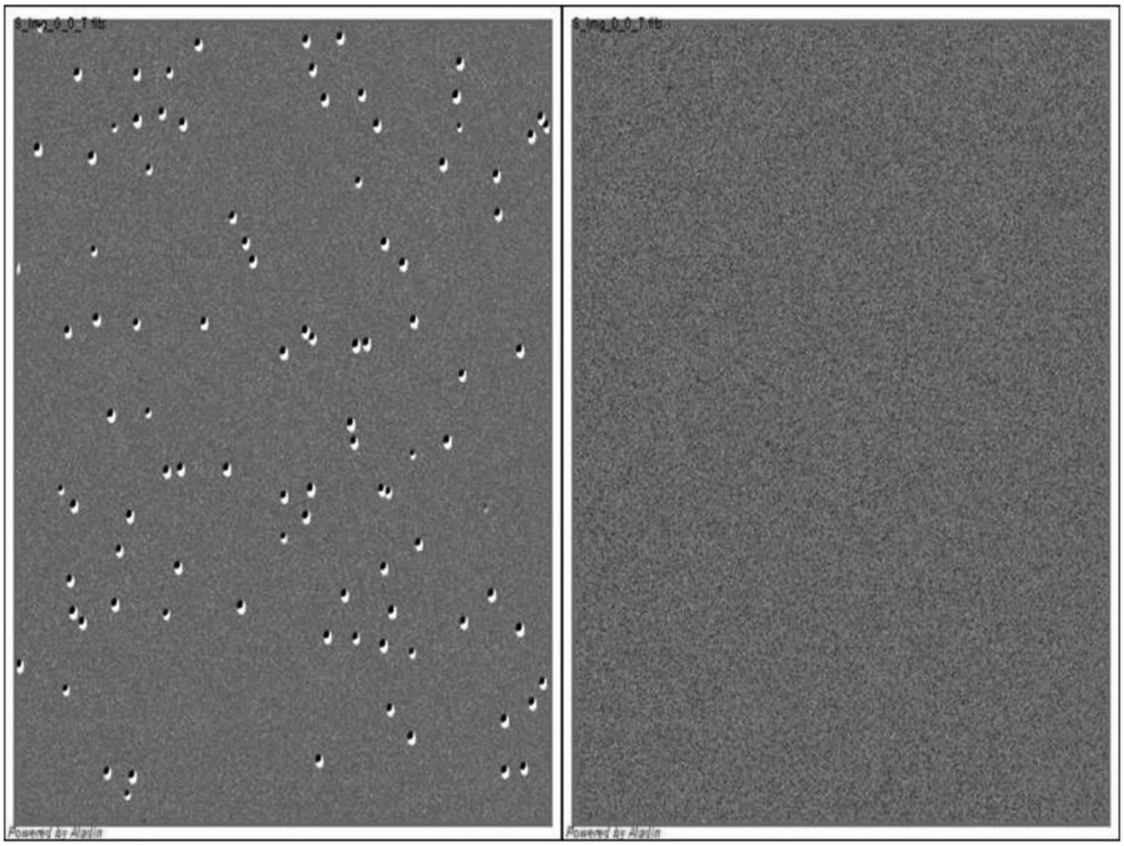

FIG. 9.—On the left is a subtraction using the OIS method and the GCB with a constant kernel. This clearly shows the systematic translation errors between the image and reference. The right subtraction is also done with the OIS method but using the DFB with a constant kernel. The systematic translation errors are corrected.

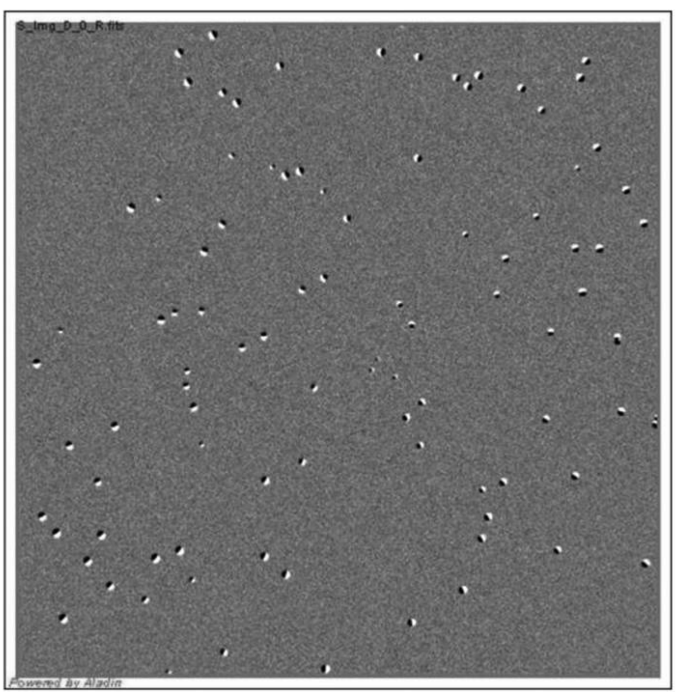

FIG. 10.—Subtraction of an image and reference rotated by 0.05°. The subtraction was done using the OIS method, with the DFB and a constant kernel.

when it is necessary to subtract across filters (i.e., I and R are taken with different filters). On occasion, this heavy-handed masking helps with the identification of moving objects in ultra-crowded star fields.

The same issue of bad pixels contaminating a subtraction arises when looking for supernovae or active galactic nuclei. The residual light from the bad pixels in the mis-subtractions make it difficult to identify either type of object, often giving a false signature (i.e., residual light that looks like the object but is not). The following procedure masks some, but not all, of these pixels (Astier 2004, private communication):

1. Use two images from the first date, I_1 and I_2
2. Use two references from the second date, R_1 and R_2
3. Subtract $S_1 = I_1 - R_1 \otimes K$ and $S_2 = I_2 - R_2 \otimes K$
4. Identify pixels on S_1 that are 3 σ above the mean sky μ but are within 1 σ on S_2
5. Identify pixels on S_2 that are 3 σ above the mean sky μ but are within 1 σ on S_1
6. Mask the identified pixels on S_1 and S_2 with the mean sky of S_1 and S_2

In other words, the hot pixels and cosmic ray hits will likely not appear on both subtractions. There will be these pixels on S_1 from I_1 and R_1 and on S_2 from I_2 and R_2 but not likely at the

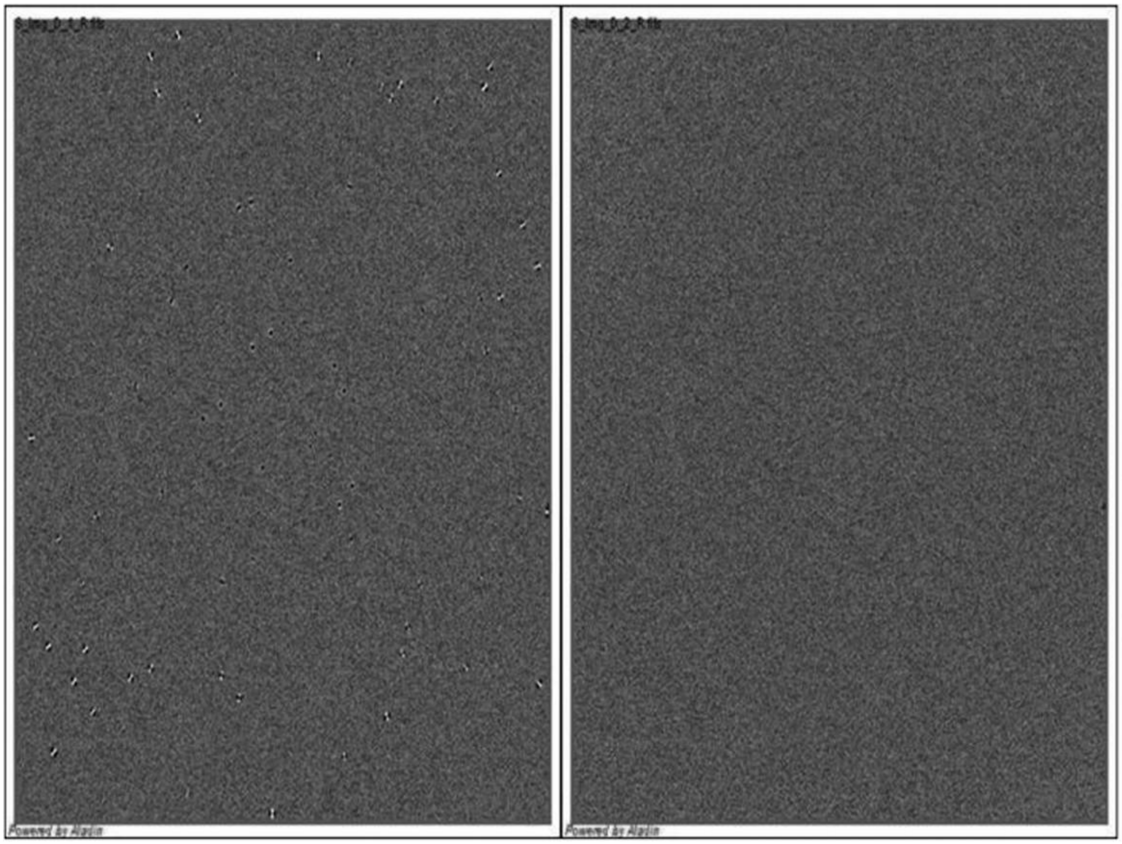

FIG. 11.—Subtraction with the OIS method using the DFB with a first degree kernel on the left and second kernel on the right. The second degree subtraction corrects the differential rotational effect.

same locations. Finding the bright pixels on S_1 but not on S_2 and on S_2 but not on S_1 identifies these two types of bad pixels, which are masked by replacing them with the mean sky of S_1 and S_2.

4. APPLICATIONS OF OIS USING IDL ALGORITHMS

Extensive testing of the IDL computer algorithms was undertaken to ensure that the OIS method as outlined in this paper was successfully implemented. The following are a few of the "show and tell" subtractions included in the validation of these algorithms.

The first is a near-Earth object (NEO) in a crowded star field shown in Figure 12. The image and reference were taken on the same evening in May 2005 as part of the Near-Earth Object Survey (LONEOS) at the Lowell Observatory (Flagstaff, AZ). These were taken along the equatorial plane of the Milky Way Galaxy

As a moving object an NEO exhibits a dark-bright signature, shown side by side (i.e., dark from the convolved reference and bright from the image). Coupled with masking of mis-subtractions (e.g., saturated stars), a moving object signature is easily detected even in a crowded star field.

Another example of a moving object is a SNe light echo. This, too, exhibits the characteristic dark-bright signature. Figure 13 shows light echoes from SN 87A in the Large Magellanic Cloud. Using the IDL computer algorithms with a first degree space-varying kernel and the delta function basis, these subtractions clearly show three rings (Rest & Suntzeff 2005). To highlight the rings, mis-subtractions (e.g., saturated stars) were masked.

The image and reference were taken by the 4-m V.M. Blanco telescope at the Cerro Tololo Inter-American Observatory over a two-year period 2002–2004. These were part of the SuperMACHO Microlensing Survey.

Objects that vary in brightness over time include supernovae (SNe). These objects exhibit a characteristic bright signature showing the residual light (i.e., the additional light contained in the convolved reference as compared to the image).

Shown in Figure 14 is SN 2006bi, a March 2006 discovery by R. Holmes, H. Devore, and the author J. P. Miller. The image and reference were separated in time by 30 days and taken with the 0.41-m telescope at the Astronomical Research Institute (Charleston, IL). SN 2006 bi was recognized as a discovery

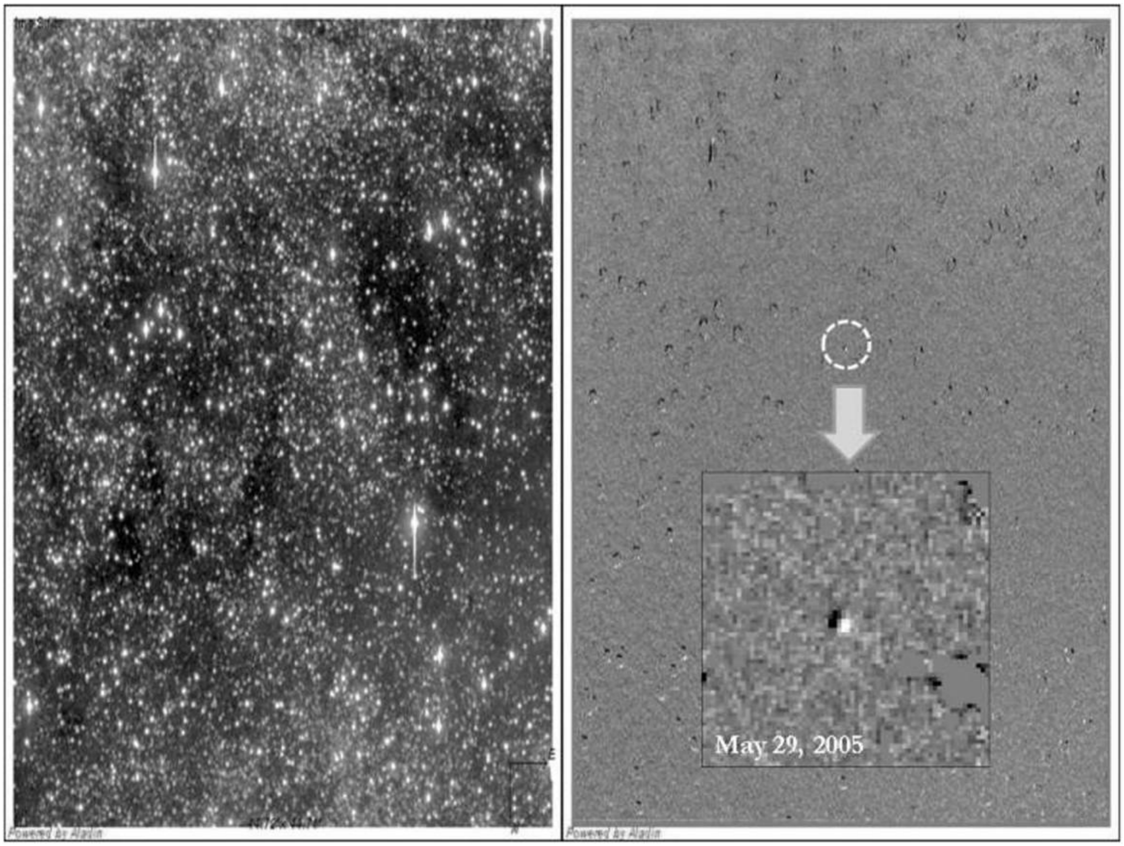

FIG. 12.—Near-Earth object in a crowded star field. Images were taken 2005 May by the Lowell Observatory as part of LONEOS. The subtraction used the first degree OIS method with the delta function basis, and included masking of saturated stars. See the electronic edition of the PASP for a color version of this figure.

by the Central Bureau for Astronomical Telegrams (CBAT; Smithsonian Astrophysical Observatory; International Astronomical Union).

5. CONCLUSIONS

The optimal image subtraction method (OIS) developed by C. Alard and R. H. Lupton provides an effective and flexible means to subtract two images. The mathematics of this method was presented, including many of the intricate details not available in the literature. Also presented were a number of issues important to the method including stamp selection and masking, alignment at the subpixel level, constant photometric flux, background correction, and masking of contaminants originating from mis-subtractions.

Computer algorithms for the OIS method were written using the Interactive Data Language (IDL). Examples of time-varying and position-varying objects were presented, which the computer algorithms successfully detected. These demonstrate the validity of the algorithms and use of the method in a wide variety of applications. These algorithms are publicly shared and will soon be available upon request. For a copy, contact the first author, J. Patrick Miller.

Dr. Pierre Astier, Laboratoire de Physique Nucléaire et de Hautes Energies (Paris), provided invaluable guidance in the development of the mathematics and algorithms used in this study. Dr. Bruce Koehn, Lowell Observatory, provided images of near-Earth objects in crowded star fields taken as part of the Near-Earth Object Survey (LONEOS). Robert E. Holmes, Jr., Astronomical Research Institute, provided images of galaxy clusters and regions along the ecliptic resulting in original discoveries of Main Belt asteroids, supernovae, and active galactic nuclei. Dr. Armin Rest, National Optical Astronomy Observatory, provided images of SN 87A taken with the 4-m Victor M. Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO, Chile). Dr. Kenneth Davis, Department of Mathematics (Hardin-Simmons University), reviewed the mathematics presented in this study. Dr. Christopher L. McNair, Dean of the Holland School of Science & Mathematics (Hardin-Simmons University), provided administrative resources and support during this study. The Aladin Sky Atlas was used to create many of the figures in this paper (Bonnarel et al. 2000).

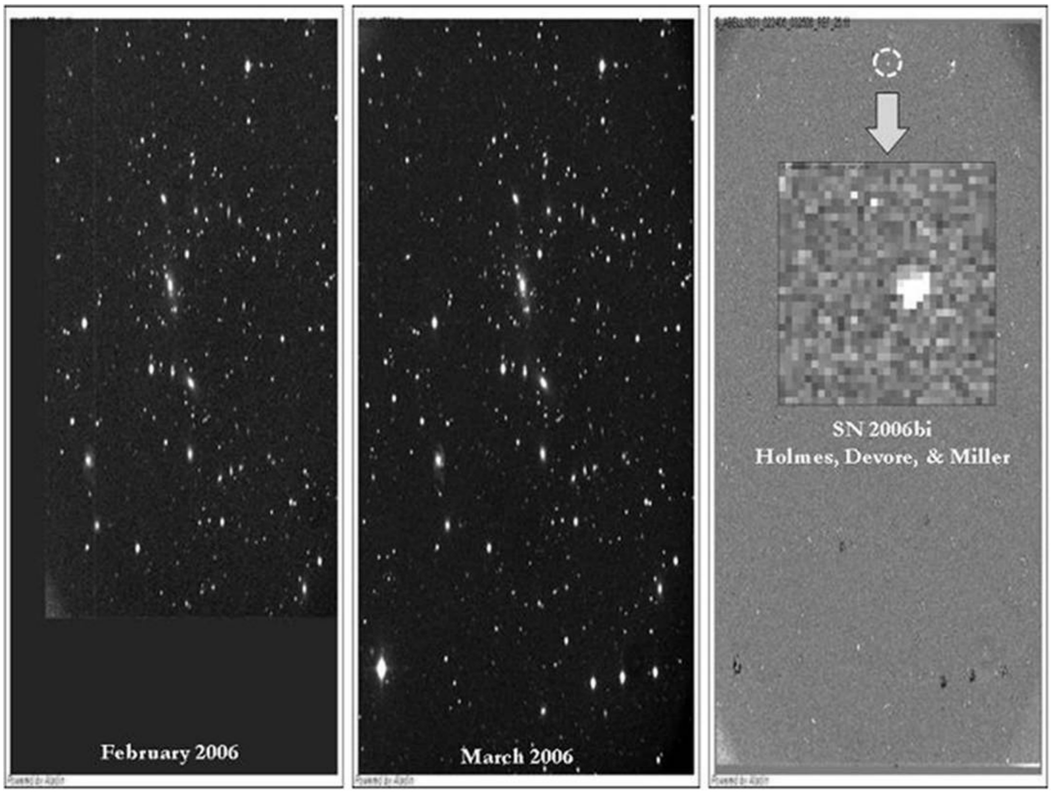

FIG. 13.—SN 87A in the Large Magellanic Cloud. The subtraction was done using the first degree OIS method and delta function basis. Clearly shown are the three rings of the light echoes. Images were taken 2002–2004 by the 4-m V.M. Blanco telescope at the Cerro Tololo Inter-American Observatory. See the electronic edition of the PASP for a color version of this figure.

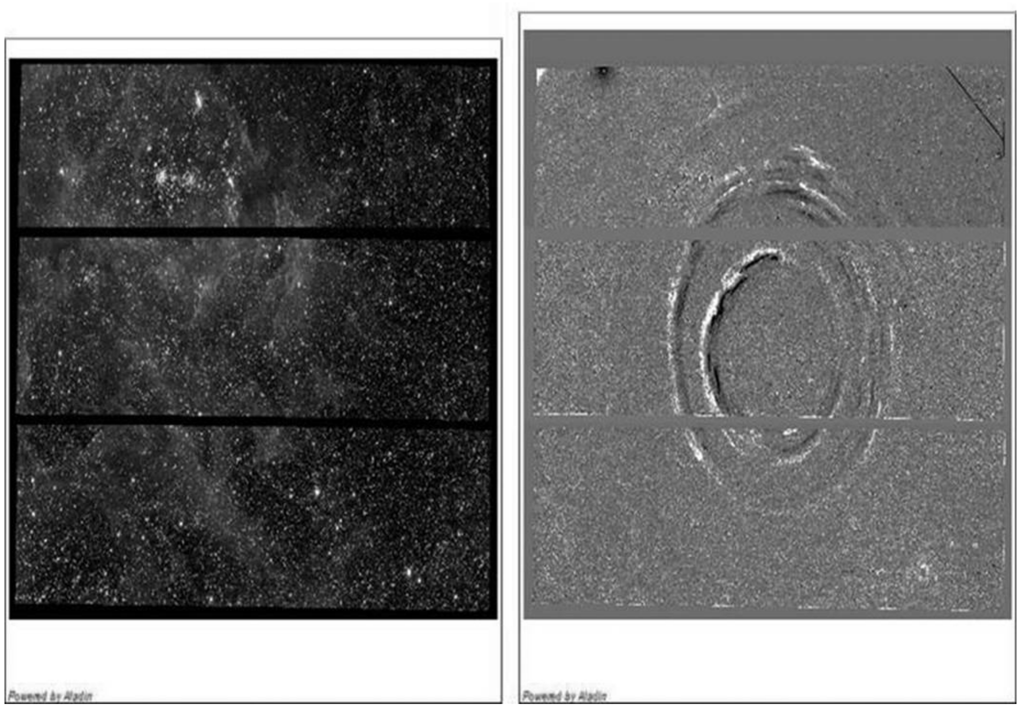

FIG. 14.—SN 2006bi in an anonymous galaxy in Abell 1831. Images were taken 2006 February-March by the Astronomical Research Institute (Charleston, IL). The subtraction used the first degree OIS method with the delta function basis.

APPENDIX A.

SPACE-VARYING KERNEL

The linear system for calculating the space-varying kernel is defined by

$$\sum_{n=1}^{N}\sum_{m=0}^{d}\sum_{l=0}^{d-m}\alpha_{n,m,l}\left[\sum_{k=1}^{P}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\hat{x}_k^{m+r}\hat{y}_k^{l+s}(R_k\otimes K_n)_{i,j}\times(R_k\otimes K_q)_{i,j}\right]=\sum_{k=1}^{P}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\hat{x}_k^{r}\hat{y}_k^{s}(I_k)_{i,j}(R_k\otimes K_q)_{i,j}, \quad \text{(A1)}$$

where $q=1,2,\ldots,N$ and for each q, $r=0,1,\ldots,d$, $s=0,1,\ldots,d-r$. For each $q=1,2,\ldots,N$ there are $\frac{1}{2}(d+1)(d+2)$ equations. This defines a $Q\times Q$ linear system where $Q=\frac{1}{2}N(d+1)(d+2)$.

Define the following two functions:

$$C_{n,m,l}(q,r,s)=\sum_{k=1}^{P}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\hat{x}_k^{m+r}\hat{y}_k^{l+s}(R_k\otimes K_n)_{i,j}(R_k\otimes K_q)_{i,j}=\sum_{k=1}^{P}\hat{x}_k^{m+r}\hat{y}_k^{l+s}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}(R_k\otimes K_n)_{i,j}(R_k\otimes K_q)_{i,j}, \quad \text{(A2)}$$

$$D(q,r,s)=\sum_{k=1}^{P}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\hat{x}_k^{r}\hat{y}_k^{s}(I_k)_{i,j}(R_k\otimes K_q)_{i,j}=\sum_{k=1}^{P}\hat{x}_k^{r}\hat{y}_k^{s}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}(I_k)_{i,j}(R_k\otimes K_q)_{i,j}. \quad \text{(A3)}$$

Let $N=2$ and $d=1$. This produces the following 6×6 linear system defined by each combination of q, r, and s. There are two basis vectors multiplied by a first-degree polynomial, which is a function of the pixel coordinates (x,y).

1.

$\mathbf{n=1,q=1}$ $\quad \boldsymbol{r=0,s=0}$

$$\alpha_{1,0,0}C_{1,0,0}(1,0,0)+\alpha_{1,0,1}C_{1,0,1}(1,0,0)+\alpha_{1,1,0}C_{1,1,0}(1,0,0)$$

$\boldsymbol{r=0,s=1}$

$$\alpha_{1,0,0}C_{1,0,0}(1,0,1)+\alpha_{1,0,1}C_{1,0,1}(1,0,1)+\alpha_{1,1,0}C_{1,1,0}(1,0,1)$$

$\boldsymbol{r=1,s=0}$

$$\alpha_{1,0,0}C_{1,0,0}(1,1,0)+\alpha_{1,0,1}C_{1,0,1}(1,1,0)+\alpha_{1,1,0}C_{1,1,0}(1,1,0) \quad \text{(A4)}$$

2.

$\mathbf{n=1,q=2}$ $\quad \boldsymbol{r=0,s=0}$

$$\alpha_{1,0,0}C_{1,0,0}(2,0,0)+\alpha_{1,0,1}C_{1,0,1}(2,0,0)+\alpha_{1,1,0}C_{1,1,0}(2,0,0)$$

$\boldsymbol{r=0,s=1}$

$$\alpha_{1,0,0}C_{1,0,0}(2,0,1)+\alpha_{1,0,1}C_{1,0,1}(2,0,1)+\alpha_{1,1,0}C_{1,1,0}(2,0,1)$$

$\boldsymbol{r=1,s=0}$

$$\alpha_{1,0,0}C_{1,0,0}(2,1,0)+\alpha_{1,0,1}C_{1,0,1}(2,1,0)+\alpha_{1,1,0}C_{1,1,0}(2,1,0) \quad \text{(A5)}$$

3.

$\mathbf{n=2,q=1}$ $\quad \boldsymbol{r=0,s=0}$

$$\alpha_{2,0,0}C_{2,0,0}(1,0,0)+\alpha_{2,0,1}C_{2,0,1}(1,0,0)+\alpha_{2,1,0}C_{2,1,0}(1,0,0)$$

$\boldsymbol{r=0,s=1}$

$$\alpha_{2,0,0}C_{2,0,0}(1,0,1)+\alpha_{2,0,1}C_{2,0,1}(1,0,1)+\alpha_{2,1,0}C_{2,1,0}(1,0,1)$$

$\boldsymbol{r=1,s=0}$

$$\alpha_{2,0,0}C_{2,0,0}(1,1,0)+\alpha_{2,0,1}C_{2,0,1}(1,1,0)+\alpha_{2,1,0}C_{2,1,0}(1,1,0) \quad \text{(A6)}$$

4.

$\mathbf{n=2,q=2}$ $\quad \boldsymbol{r=0,s=0}$

$$\alpha_{2,0,0}C_{2,0,0}(2,0,0)+\alpha_{2,0,1}C_{2,0,1}(2,0,0)+\alpha_{2,1,0}C_{2,1,0}(2,0,0)$$

$\boldsymbol{r=0,s=1}$

$$\alpha_{2,0,0}C_{2,0,0}(2,0,1)+\alpha_{2,0,1}C_{2,0,1}(2,0,1)+\alpha_{2,1,0}C_{2,1,0}(2,0,1)$$

$\boldsymbol{r=1,s=0}$

$$\alpha_{2,0,0}C_{2,0,0}(2,1,0)+\alpha_{2,0,1}C_{2,0,1}(2,1,0)+\alpha_{2,1,0}C_{2,1,0}(2,1,0) \quad \text{(A7)}$$

Equation for $\mathbf{q=1}$, $\mathbf{r=0}$, $\mathbf{s=0}$

$$\alpha_{1,0,0}C_{1,0,0}(1,0,0)+\alpha_{1,0,1}C_{1,0,1}(1,0,0)+\alpha_{1,1,0}C_{1,1,0}(1,0,0)+\alpha_{2,0,0}C_{2,0,0}(1,0,0)+\alpha_{2,0,1}C_{2,0,1}(1,0,0)+\alpha_{2,1,0}C_{2,1,0}(1,0,0)=D(1,0,0) \quad \text{(A8)}$$

Equation for $\mathbf{q=1}$, $\mathbf{r=0}$, $\mathbf{s=1}$

$$\alpha_{1,0,0}C_{1,0,0}(1,0,1)+\alpha_{1,0,1}C_{1,0,1}(1,0,1)+\alpha_{1,1,0}C_{1,1,0}(1,0,1)+\alpha_{2,0,0}C_{2,0,0}(1,0,1)+\alpha_{2,0,1}C_{2,0,1}(1,0,1)+\alpha_{2,1,0}C_{2,1,0}(1,0,1)=D(1,0,1) \quad \text{(A9)}$$

Equation for $\mathbf{q = 1}$, $\mathbf{r = 1}$, $\mathbf{s = 0}$

$$\alpha_{1,0,0}C_{1,0,0}(1,1,0) + \alpha_{1,0,1}C_{1,0,1}(1,1,0) + \alpha_{1,1,0}C_{1,1,0}(1,1,0) + \alpha_{2,0,0}C_{2,0,0}(1,1,0) + \alpha_{2,0,1}C_{2,0,1}(1,1,0) + \alpha_{2,1,0}C_{2,1,0}(1,1,0) = D(1,1,0) \tag{A10}$$

Equation for $\mathbf{q = 2}$, $\mathbf{r = 0}$, $\mathbf{s = 0}$

$$\alpha_{1,0,0}C_{1,0,0}(2,0,0) + \alpha_{1,0,1}C_{1,0,1}(2,0,0) + \alpha_{1,1,0}C_{1,1,0}(2,0,0) + \alpha_{2,0,0}C_{2,0,0}(2,0,0) + \alpha_{2,0,1}C_{2,0,1}(2,0,0) + \alpha_{2,1,0}C_{2,1,0}(2,0,0) = D(2,0,0) \tag{A11}$$

Equation for $\mathbf{q = 2}$, $\mathbf{r = 0}$, $\mathbf{s = 1}$

$$\alpha_{1,0,0}C_{1,0,0}(2,0,1) + \alpha_{1,0,1}C_{1,0,1}(2,0,1) + \alpha_{1,1,0}C_{1,1,0}(2,0,1) + \alpha_{2,0,0}C_{2,0,0}(2,0,1) + \alpha_{2,0,1}C_{2,0,1}(2,0,1) + \alpha_{2,1,0}C_{2,1,0}(2,0,1) = D(2,0,1) \tag{A12}$$

Equation for $\mathbf{q = 2}$, $\mathbf{r = 1}$, $\mathbf{s = 0}$

$$\alpha_{1,0,0}C_{1,0,0}(2,1,0) + \alpha_{1,0,1}C_{1,0,1}(2,1,0) + \alpha_{1,1,0}C_{1,1,0}(2,1,0) + \alpha_{2,0,0}C_{2,0,0}(2,1,0) + \alpha_{2,0,1}C_{2,0,1}(2,1,0) + \alpha_{2,1,0}C_{2,1,0}(2,1,0) = D(2,0,1) \tag{A13}$$

This 6×6 linear system is solved for the vector: $[\alpha_{1,0,0}, \alpha_{1,0,1}, \alpha_{1,1,0}; \alpha_{2,0,0}, \alpha_{2,0,1}, \alpha_{2,1,0}]$ where the first set of three components define the coefficients of the polynomial factor associated with the first basis vector and the second set of three define the coefficients of the polynomial factor associated with the second basis vector.

APPENDIX B

DIFFERENTIAL BACKGROUND

Let I be an image with dimensions of $S_R \times S_C$ having a background sky of mean μ and standard deviation σ. Define the set $B = \{I_{i,j} : \ |I_{i,j} - \mu| \leq \sigma\}$ as the set of pixels in the image J within 1 σ of the mean sky, and let Λ be the set of indices (i.e., pixel coordinates) of the elements of B. B is the background not appearing in any light sources.

Using a least-squares calculation to find the values of $c_{k,l}$ over the set of background pixels B, the error $E_{i,j} = I_{i,j} - p(i,j)$ where the pixel coordinates $(i,j) \in \Lambda$, and the sum of the squares of the errors is given by

$$F = \sum_{(i,j)\in\Lambda} E_{i,j}^2 = \sum_{(i,j)\in\Lambda} [I_{i,j} - p(i,j)]^2 = \sum_{(i,j)\in\Lambda} \left[I_{i,j} - \sum_{k=0}^{\hat{d}} \sum_{l=0}^{\hat{d}-k} c_{k,l} i^k j^l \right]^2. \tag{B1}$$

The least-squares solution is given by $\nabla F = 0$. This produces the $Z \times Z$ linear system where $Z = \frac{1}{2}(\hat{d}+1)(\hat{d}+2)$:

$$\sum_{k=0}^{\hat{d}} \sum_{l=0}^{\hat{d}-1} c_{k,l} \sum_{(i,j)\in\Lambda} i^{k+n} j^{l+m} = \sum_{(i,j)\in\Lambda} I_{i,j} i^n j^m, \tag{B2}$$

where $0 \leq n \leq \hat{d}, 0 \leq m \leq \hat{d} - n$.

Suppose a second degree polynomial is sought for the background correction. In this case, $\hat{d} = 2$, which gives a 6×6 linear system. The following is that linear system. Note for each n and m, one row of that system is produced.

Row no. 1: $n = 0$, $m = 0$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^0 j^0 + c_{0,1} \sum_{(i,j)\in\Lambda} i^0 j^1 + c_{0,2} \sum_{(i,j)\in\Lambda} i^0 j^2 + c_{1,0} \sum_{(i,j)\in\Lambda} i^1 j^0 + c_{1,1} \sum_{(i,j)\in\Lambda} i^1 j^1 + c_{2,0} \sum_{(i,j)\in\Lambda} i^2 j^0 = \sum_{(i,j)\in\Lambda} I_{i,j} i^0 j^0 \tag{B3}$$

Row no. 2: $n = 0$, $m = 1$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^0 j^1 + c_{0,1} \sum_{(i,j)\in\Lambda} i^0 j^2 + c_{0,2} \sum_{(i,j)\in\Lambda} i^0 j^3 + c_{1,0} \sum_{(i,j)\Lambda} i^1 j^1 + c_{1,1} \sum_{(i,j)\in\Lambda} i^1 j^2 + c_{2,0} \sum_{(i,j)\in\Lambda} i^2 j^1 = \sum_{(i,j)\in\Lambda} I_{i,j} i^0 j^1 \tag{B4}$$

Row no. 3: $n = 0$, $m = 2$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^0 j^2 + c_{0,1} \sum_{(i,j)\in\Lambda} i^0 j^3 + c_{0,2} \sum_{(i,j)\in\Lambda} i^0 j^4 + c_{1,0} \sum_{(i,j)\in\Lambda} i^1 j^2 + c_{1,1} \sum_{(i,j)\in\Lambda} i^1 j^3 + c_{2,0} \sum_{(i,j)\in\Lambda} i^2 j^2 = \sum_{(i,j)\in\Lambda} I_{i,j} i^0 j^2 \tag{B5}$$

Row no. 4: $n = 1$, $m = 0$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^1 j^0 + c_{0,1} \sum_{(i,j)\in\Lambda} i^1 j^1 + c_{0,2} \sum_{(i,j)\in\Lambda} i^1 j^2 + c_{1,0} \sum_{(i,j)\in\Lambda} i^2 j^0 + c_{1,1} \sum_{(i,j)\in\Lambda} i^2 j^1 + c_{2,0} \sum_{(i,j)\in\Lambda} i^3 j^0 = \sum_{(i,j)\in\Lambda} I_{i,j} i^1 j^0 \tag{B6}$$

Row no. 5: $n = 1$, $m = 1$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^1 j^1 + c_{0,1} \sum_{(i,j)\in\Lambda} i^1 j^2 + c_{0,2} \sum_{(i,j)\in\Lambda} i^1 j^3 + c_{1,0} \sum_{(i,j)\in\Lambda} i^2 j^1 + c_{1,1} \sum_{(i,j)\in\Lambda} i^2 j^2 + c_{2,0} \sum_{(i,j)\in\Lambda} i^3 j^1 = \sum_{(i,j)\in\Lambda} I_{i,j} i^1 j^1 \tag{B7}$$

Row no. 6: $n = 2$, $m = 0$

$$c_{0,0} \sum_{(i,j)\in\Lambda} i^2 j^0 + c_{0,1} \sum_{(i,j)\in\Lambda} i^2 j^1 + c_{0,2} \sum_{(i,j)\in\Lambda} i^2 j^2 + c_{1,0} \sum_{(i,j)\in\Lambda} i^3 j^0 + c_{1,1} \sum_{(i,j)\in\Lambda} i^3 j^1 + c_{2,0} \sum_{(i,j)\in\Lambda} i^4 j^0 = \sum_{(i,j)\in\Lambda} I_{i,j} i^2 j^0 \tag{B8}$$

APPENDIX C

CONSTANT PHOTOMETRIC RATIO

The following is a proof that a redefinition of the basis vectors produces a kernel with an integral of 1 needed for a constant photometric ratio. As shown in § 2.5 the original set of basis vectors $\{K_n :\ n = 1, 2, \ldots, N\}$ are redefined as a new set of basis vectors $\{\hat{K}_n :\ n = 1, 2, \ldots, N\}$ such that

$$\|\hat{K}_n\| = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

for $n = 1, 2, \ldots, N$. This gives the following for the integral of the kernel $\|K\|$:

$$\begin{aligned} \|K\| &= \sum_{l=1}^{L} \sum_{m=1}^{L} \hat{K}_{l,m}(x,y) \\ &= \sum_{l=1}^{L} \sum_{m=1}^{L} (\hat{K}_1)_{l,m} + \sum_{l=1}^{L} \sum_{m=1}^{L} \sum_{n=2}^{N} a_n(x,y) (\hat{K}_n)_{l,m}, \end{aligned} \tag{C1}$$

$$\begin{aligned} \|K\| &= \sum_{l=1}^{L} \sum_{m=1}^{L} \hat{K}_{l,m}(x,y) \\ &= \sum_{l=1}^{L} \sum_{m=1}^{L} (\hat{K}_1)_{l,m} + \sum_{n=2}^{N} a_n(x,y) \sum_{l=1}^{L} \sum_{m=1}^{L} (\hat{K}_n)_{l,m}, \end{aligned} \tag{C2}$$

$$\|K\| = \|\hat{K}_1\| + \sum_{n=2}^{N} a_n(x,y) \|\hat{K}_n\| = 1. \tag{C3}$$

APPENDIX D

SUBPIXEL TRANSLATIONS

Let I and R be $S_R \times S_C$ images, I is an arbitrary image and R is the reference image in a subtraction. Let K be an $L \times L$ kernel, where L is odd and $L = 2\Delta w + 1$.

The convolution of the arbitrary image I, $I \otimes K$, is an $S_R \times S_C$ image such that:

$$(I \otimes K)_{i,j} = \sum_{u=i-\Delta w}^{i+\Delta w} \sum_{v=j-\Delta w}^{j+\Delta w} I_{u,v} K_{u+1-i+\Delta w, v+1-j+\Delta w}. \tag{D1}$$

Let K_H be a horizontal subpixel shift kernel and K_V be a vertical subpixel shift kernel:

$$\begin{aligned} &(K_H)_{i,\Delta w+1} \neq 0 \quad \text{for } i = 1, \ldots, L \quad \text{and } 0, \quad \text{otherwise} \\ &(K_V)_{\Delta w+1,j} \neq 0 \quad \text{for } j = 1, \ldots, L \quad \text{and } 0, \quad \text{otherwise.} \end{aligned} \tag{D2}$$

This gives the following for $I \otimes K_H$ and $I \otimes K_V$, where I is an arbitrary image:

$$\begin{aligned} (I \otimes K_H)_{i,j} &= \sum_{u=i-\Delta w}^{i+\Delta w} I_{u,j} (K_H)_{u+1-i+\Delta w, \Delta w+1}, \\ (I \otimes K_V)_{i,j} &= \sum_{v=j-\Delta w}^{j+\Delta w} I_{i,v} (K_V)_{\Delta w+1, v+1-j+\Delta w}. \end{aligned} \tag{D3}$$

Consider the convolution of the convolution $(R \otimes K_{H1}) \otimes K_V$, a horizontal subpixel shift followed by a vertical subpixel shift, and $(R \otimes K_V) \otimes K_H$, a vertical subpixel shift followed by a horizontal subpixel shift:

$$[(R \otimes K_H) \otimes K_V]_{i,j} = \sum_{v=j-\Delta w}^{j+\Delta w} (R \otimes K_H)_{i,v} (K_V)_{\Delta w+1, v+1-j+\Delta w}$$
$$= \sum_{v=j-\Delta w}^{j+\Delta w} \left[\sum_{u=i-\Delta w}^{i+\Delta w} R_{u,v} (K_H)_{u+1-i+\Delta w, \Delta w+1} \right] \times (K_V)_{\Delta w+1, v+1-j+\Delta w}, \quad \text{(D4)}$$

$$[(R \otimes K_H) \otimes K_V]_{i,j} = \sum_{v=j-\Delta w}^{j+\Delta w} \sum_{u=i-\Delta w}^{i+\Delta w} R_{u,v} (K_H)_{u+1-i+\Delta w, \Delta w+1} \times (K_V)_{\Delta w+1, v+1-j+\Delta w}, \quad \text{(D5)}$$

$$[(R \otimes K_H) \otimes K_V]_{i,j} = \sum_{u=i-\Delta w}^{i+\Delta w} \sum_{v=j-\Delta w}^{j+\Delta w} R_{u,v} (K_H)_{u+1-i+\Delta w, \Delta w+1} \times (K_V)_{\Delta w+1, v+1-j+\Delta w}, \quad \text{(D6)}$$

$$[(R \otimes K_H) \otimes K_V]_{i,j} = \sum_{u=i-\Delta w}^{i+\Delta w} \left[\sum_{v=j-\Delta w}^{j+\Delta w} R_{u,v} (K_V)_{\Delta w+1, v+1-j+\Delta w} \right] \times (K_H)_{u+1-i+\Delta w, \Delta w+1}, \quad \text{(D7)}$$

$$[(R \otimes K_H) \otimes K_V]_{i,j} = \sum_{u=i-\Delta w}^{i+\Delta w} (R \otimes K_V)_{u,j} (K_H)_{u+1-i+\Delta w, \Delta w+1}$$
$$= [(R \otimes K_V) \otimes K_H]_{i,j}. \quad \text{(D8)}$$

Therefore, when K_H and K_V are subpixel shift kernels (horizontal and vertical) the convolution of the convolution is commutative.

Furthermore, $\|K_H\| = \|K_V\| = 1$ (i.e., conserve the photometric flux). The following is the proof for the second degree interpolating polynomial producing a horizontal subpixel shift; however, the proof for both a horizontal and vertical shift of any degree follows the same approach:

$$K_H = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2}(u-i) \times (u-i-1) & -(u-i+1) \times (u-i-1) & \frac{1}{2}(u-i+1) \times (u-i) \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(D9)}$$

$$\|K_H\| = \tfrac{1}{2}(u-i)(u-i-1) - (u-i+1)(u-i-1) + \tfrac{1}{2}(u-i+1)(u-i) \quad \text{(D10)}$$

$$\|K_H\| = \tfrac{1}{2}(u-i)^2 - \tfrac{1}{2}(u-i) - (u-i)^2 + 1 + \tfrac{1}{2}(u-i)^2 + \tfrac{1}{2}(u-i) = 1 \quad \text{(D11)}$$

REFERENCES

Alard, C. 2000, A&AS, 144, 363

Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325

Bond, I. A., Abe, F., & Dodd, R. J.et al. 2001, MNRAS, 327, 868

Bonnarel, B., et al. 2000, A&AS, 143, 33

Irwin, J., & Irwin, M. 2003, http://www.ast.cam.ac.uk/vdfs/docs/reports/subtract2/

Israel, H., Hessman, F. V., & Schuh, S. 2007, AN, 328, 16

Landsman, W. 2004, IDL Astronomy User's Library, GSFC, http://idlastro.gsfc.nasa.gov/

Miller, J. P. 2007, Doctoral Thesis

Phillips, A. C., & Davis, L. E. 1995, ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 297

Rest, A., & Suntzeff, N.et al. 2005, Nature, 438, 1132

Tomaney, A. B., Crotts, A. P. S. 1996, AJ, 112, 2872