WPHOT Solution for Flux, Position, and Proper Motion

Document number: WSDC-D-T041

1. Overview of Previous WPHOT Versions

The v1-v5 versions of the WPHOT module solved for point-source fluxes and positions by applying the procedure described in the WPHOT SDS (WSDC D-D006), section 3.1.1. Sources are initially detected in multi-band co-added images in the MDET module and passed to WPHOT for extraction by fitting PSF templates to each frame in the co-add stack at each source position. The detections are processed in order of descending brightness, and as each is processed, nearby sources are identified and grouped into an ensemble of potentially blended sources. The source about which other sources are grouped is called the *primary component* of the ensemble. The idea is to include the entire ensemble in the fit in order to get the best estimates of the primary component's parameters; after that has been accomplished, the primary component is output, and all other components' solutions are discarded, since these sources will each come up later in the role of a primary component. Often the ensemble consists only of a primary component, particularly in regions of low source density. When a source is output, its flux-normalized PSF is subtracted from all frames in the stack.

Since prior uncertainties for frame-pixel fluxes are available, it is possible to compute a chi-square figure of merit for any solution for a set of blended sources. The fitting is done by maximizing the conditional probability of the observations given the fitting parameters. Since the errors are all modeled as Gaussian (which may include Poisson processes in their Gaussian asymptotic limits), this is equivalent to searching for the chi-square minimum in an n_p-dimensional space, where n_p is the number of parameters in the model, i.e., the sum over the blend-component ensemble of the number of available bands N_λ plus $2N_B$, the latter being the RA and Dec coordinates of the N_B blend components.

To summarize the SDS, we have an observational model

$$\rho_{\lambda i} = b_{\lambda i} + \nu_{\lambda i} + \sum_{n=1}^{N_B} \left(f_{\lambda n} H_\lambda (\vec{r}_{\lambda i} - \vec{s}_n) \right)$$

where $\rho_{\lambda i}$ is the flux in the i^{th} usable observed pixel at the position given by the 2-vector $\vec{r}_{\lambda i}$ in the wavelength band λ, $\vec{s}_n$ is a 2-vector giving the position of the n^{th} blend component, $f_{\lambda n}$ is the flux of the n^{th} component, $H_\lambda(\vec{r})$ is the PSF, $b_{\lambda i}$ is the local background estimated in an annulus around pixel

i, and $v_{\lambda i}$ is the noise in the i^{th} pixel, with prior noise variance $\sigma^2_{i\lambda}$, which includes uncertainty due to the PSF model based on an initial flux estimate from aperture photometry. For details on the computation of the background and noise, the reader is referred to the SDS.

The reduced chi-square figure of merit is defined as

$$\chi^2 = \frac{1}{N_{obs} - n_p} \sum_{\lambda} \sum_{i} \left(\frac{\rho_{\lambda i} - b_{\lambda i} - \sum_{n=1}^{N_B} \hat{f}_{\lambda n} H_{\lambda}(\vec{r}_{\lambda i} - \hat{\vec{s}}_n)}{\sigma_{\lambda i}} \right)^2$$

where N_{obs} is the total number of usable pixels, and the circumflexes on $f_{\lambda n}$ and $\vec{s}_n$ indicate that current estimates of these values are used to evaluate chi-square. These are the parameters to be estimated via minimization of chi-square, and hence the solution depends on finding the minimum in the n_p-dimensional parameter space. The solution proceeds as in all nonlinear fitting problems: zeroth-order estimates are used to evaluate the function and its derivatives with respect to the fitting parameters, and then an extremum-finding algorithm is applied. Since chi-square is to be minimized, we seek a minimum, and the algorithm of choice is the *gradient descent method*, also known as the *method of steepest descent* (not to be confused with the integral-approximation algorithm with the same name).

Once the minimum chi-square has been found, the primary component is output and the other solutions are discarded. The only exception to this involves active deblending. The source ensemble starts out composed of a primary component and nearby sources that were detected on the co-adds. The solution for the ensemble is called *passive deblending*. If the chi-square for that solution is too high, an attempt is made (subject to controllable thresholds) to find another source in the neighborhood that was not detected on the co-adds but is blended with the other ensemble sources. We will leave the details of this processing to the SDS, but the bottom line is that if the ensemble reduced chi-square can be made smaller (by another controllable threshold) by including the new source, then it is kept and called an *actively deblended* source. Since this source is not in the list of co-add detections, it will not come up later as a primary component, and so it is output immediately after the primary component. It is theoretically possible to have more than one actively deblended component.

2. Addition of Proper Motion Estimation

The change made to the observational model in the v6 version of WPHOT is to replace $\vec{s}_n$ for the primary component with a frame-dependent position that is a linear function of time. Since the primary component is usually component number 1 in the ensemble, for notational convenience we will just assign it $n = 1$ herein. Then

$$\vec{s}_1 \Rightarrow \vec{s}_{1i} = \vec{s}_{10} + \vec{\mu}(t_i - t_0)$$

where $\vec{s}_{10}$ is the position of the primary component at a fiducial time t_0, $\vec{\mu}$ is the 2-vector proper

motion of the primary component, and t_i is the time tag for the frame in the co-add stack that contains pixel *i*. The initial estimates for the components of $\vec{\mu}$ are zero. These will evolve to non-zero values if this reduces chi-square. The influence of the proper motion on chi-square is through its displacements of the PSF in the frames of the co-add stack. So the expression for the reduced chi-square becomes

$$\chi^2 = \frac{1}{N_{obs} - n_p} \sum_{\lambda} \sum_{i} \left(\frac{\rho_{\lambda i} - b_{\lambda i} - \hat{f}_{\lambda 1} H_\lambda (\vec{r}_{\lambda i} - \hat{\vec{s}}_{1i} - \hat{\vec{\mu}}(t_i - t_0)) - \sum_{n=2}^{N_B} \hat{f}_{\lambda n} H_\lambda (\vec{r}_{\lambda i} - \hat{\vec{s}}_n)}{\sigma_{\lambda i}} \right)^2$$

The inclusion of the 2-vector $\vec{\mu}$ enlarges n_p by 2. The rest of the processing is the same as before except for the larger parameter space and the additional output source parameters for the proper motion components and their associated uncertainties. Since actively deblended sources never take on the role of primary component of an ensemble, *their proper motion is not estimated and will be null in the final products*.

The uncertainty estimates for proper motion are obtained in the same manner as the parameter uncertainties of previous versions of WPHOT. As described in the SDS, the Fisher information matrix *G* is computed, and the error covariance matrix γ expressing the parameter uncertainties is the inverse of *G*:

$$\ln P(\vec{\rho}|\vec{z}, N_B) = -\frac{1}{2} \sum_{\lambda} \sum_{i} \left(\frac{\rho_{\lambda i} - b_{\lambda i} - f_{\lambda 1} H_\lambda (\vec{r}_{\lambda i} - \vec{s}_{1i} - \vec{\mu}(t_i - t_0)) - \sum_{n=2}^{N_B} f_{\lambda n} H_\lambda (\vec{r}_{\lambda i} - \vec{s}_n)}{\sigma_{\lambda i}} \right)^2 + \text{const.}$$

$$G \equiv -\left\langle \frac{\partial}{\partial \vec{z}} \frac{\partial^T}{\partial \vec{z}} \ln P(\vec{\rho}|\vec{z}, N_B) \right\rangle$$

$$\gamma = G^{-1}$$

$$\sigma(z_j)^2 = \gamma_{jj}$$

where $\vec{\rho}$ is the vector of all pixels used for the calculation, with N_{obs} elements, and $\vec{z}$ is the vector of all parameters being fit, with n_p elements. The *G* and γ matrices are therefore $n_p \times n_p$ in size. The components of the $\vec{z}$ vector are [$\vec{s}_{10}$, ($f_{\lambda 1}$: λ = 1 to N_λ), { $\vec{s}_n$, ($f_{\lambda n}$: λ = 1 to N_λ)} : *n* = 2 to N_B, $\vec{\mu}$].

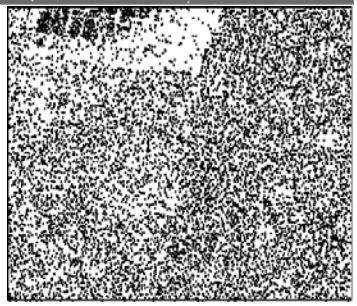
The dependence of γ on $\vec{\mu}$ is via the dependence of H_λ on $\vec{s}_1$, whose derivatives may take on positive, negative, or zero values, depending on where in the PSF a given pixel falls, and since the PSF models are numerical, not algebraic, and also since the total dependence involves a sum over all pixels used in the computation, how the new proper motion parameters affect the uncertainties of the fluxes is most readily investigated numerically rather than algebraically. Testing of the new algorithm will therefore include inspection of all elements of γ and comparison to those of the

previous algorithm in order to analyze the effect of including proper motion on estimated flux, especially for the case in which proper motion is negligibly small. The first results of such testing are described in the next section.

The value of the fiducial time t_0 used in initial testing has been JD 2455197.5 (January 1.0 2010). Since the fiducial time is the epoch of the nominal positions of sources with significant proper motion, it is recommended to use JD 2455400.0 instead, which is essentially the midpoint of the full survey period.

3. Testing

Initial testing has been done on the tile used for the AllWISE proposal 0626p151. This tile contains the Y dwarf WISE J041022.71+150248.4. The coadd is 1.6×1.6 degrees, centered at RA = 62.6 degrees and Dec = 15.0 degrees. Only W1 and W2 frames were used. Coadd detection produced 10,251 sources. None were added via active deblending, which was turned off for this test case. Only 14 cases of passive deblending occurred. 211 Level 1b frames per band were used. The earliest and latest frame DATE_OBS values are 2010-02-15T05:24:28.902 and 2010-08-27T00:44:10.649, respectively, so the time span is about 192 days, 18 hours, and 20 minutes and includes both four-band-cryo and three-band-cryo data. This case was processed with both the v6 and v5 versions of WPHOT (the v5 version is identical to v6 except for the absence of proper motion estimation). The W1 and W2 coadds are shown below on the left and right, log-99% scale and linear-99%-scale, respectively.



The W2 data obviously suffer from the accidental inclusion of some frames contaminated by scattered moonlight. The pattern of source detection is shown on the right and clearly reflects the bad-frame pattern. The tests described in this section will be repeated with a cleaner test case as soon as possible. For now, we assume that these shortcomings may constitute a stress test that probably has little effect on the mechanics of estimating source positions, fluxes, and proper

motions other than possibly to degrade it somewhat.

It is reasonable to expect that sources with significant proper motion, hence blurring in the coadds and imperfect alignment in the individual-scan frames, would suffer extra flux estimation error that would be removed by including proper motion in the model. But some interest has been expressed in whether the addition of proper motion to the set of estimated parameters has any negative impact on the flux estimation when the true proper motion is negligible. By negative impact, we mean both additional error in the flux and enlargement of the flux uncertainty. This question was investigated two ways: (a.) a direct comparison of the fluxes and uncertainties in the v6 results against the v5 results from the exact same input; (b.) examination of the formal error correlation matrix for the v6 version. The extent to which flux estimation is improved when there is significant proper motion will require much more extensive analysis and hence will be postponed for now.

One probe of the effect of proper motion inclusion on flux and flux uncertainty is the ratio of fluxes measured by the v5 and v6 versions of WPHOT. We define

$$R_{W1} \equiv \left\langle \frac{W1_{v6}}{W1_{v5}} \right\rangle$$

$$R_{W2} \equiv \left\langle \frac{W2_{v6}}{W2_{v5}} \right\rangle$$

$$R_{\sigma 1} \equiv \left\langle \frac{\sigma(W1_{v6})}{\sigma(W1_{v5})} \right\rangle$$

$$R_{\sigma 2} \equiv \left\langle \frac{\sigma(W2_{v6})}{\sigma(W2_{v5})} \right\rangle$$

$$R_{\chi^2_{W1}} \equiv \left\langle \frac{\chi^2(W1_{v6})}{\chi^2(W1_{v5})} \right\rangle$$

$$R_{\chi^2_{W2}} \equiv \left\langle \frac{\chi^2(W2_{v6})}{\chi^2(W2_{v5})} \right\rangle$$

$$R_{\chi^2} \equiv \left\langle \frac{\chi^2_{v6}}{\chi^2_{v5}} \right\rangle$$

where the fluxes W1 and W2 are measured in DN, the chi-squares are all reduced chi-squares (i.e., chi-square divided by the number of degrees of freedom), the first two of which are for the fluxes in the bands indicated, and the last is the overall chi-square, the one that the algorithm minimizes, which for v6 has the extra two degrees of freedom contributed by the proper motion (given identical deblending activity in the v5 and v6 processing of the test case). The averaging was carried out over the detections processed by v5 and v6.

The samples averaged to produce these statistics depended somewhat on flux level. Very near zero, large differences were seen, which is to be expected. To prevent this from obscuring the comparison,

several constraints on flux were tried. When the only constraint was that W1 and W2 fluxes be greater than zero in both v5 and v6, the average flux ratios were slightly greater than 1, with values of about 1.0326 and 1.0105 in W1 and W2 respectively. When the fluxes were all required to be greater than 3-sigma, these values dropped to 0.9885 and 1.0020, respectively. Several tests were done eliminating the brightest sources, but these had a negligible effect. The ratio of flux uncertainties and chi-squares depended less on brightness. More work is planned to investigate brightness dependences of the samples averaged to obtain these parameters. The statistical results are summarized in the table below.

	W1,W2 > 0	W1,W2 > 3σ	10^5 < W1,W2 > 3σ	10^4 < W1,W2 > 3σ
R_{W1}	1.0326415	0.9884621	0.9884359	0.988237
R_{W2}	1.0105182	1.0020322	1.0020549	1.0021055
$R_{\sigma 1}$	0.9970112	0.9968084	0.9967994	0.9966804
$R_{\sigma 2}$	0.9984549	0.9981387	0.9980943	0.9980355
$R_{\chi^2_{W1}}$	1.0583574	1.0602149	1.0601076	1.0596383
$R_{\chi^2_{W2}}$	1.0071218	1.0072376	1.0069995	1.0056945
R_{χ^2}	1.0325043	1.0333724	1.0332161	1.0323538
Npts	9693	9286	9254	9092

The standard deviations about these mean values are:

	W1,W2 > 0	W1,W2 > 3σ	10^5 < W1,W2 > 3σ	10^4 < W1,W2 > 3σ
$\sigma(R_{W1})$	3.8393898	0.0411089	0.0411707	0.0415036
$\sigma(R_{W2})$	0.1856340	0.0432361	0.0433075	0.0436874
$\sigma(R_{\sigma 1})$	0.0466752	0.0330567	0.0330860	0.0333239
$\sigma(R_{\sigma 2})$	0.0404073	0.0249154	0.0246784	0.0248486
$\sigma(R_{\chi^2_{W1}})$	0.1581026	0.1606310	0.1607951	0.1618636
$\sigma(R_{\chi^2_{W2}})$	0.0429418	0.0354310	0.0345969	0.0318948
$\sigma(R_{\chi^2})$	0.0757564	0.0732473	0.0730948	0.0728373

The error correlation matrix was computed from the formal error covariance matrix and averaged over all but nine unblended sources. For reasons not yet known but which will be investigated and which are not peculiar to v6, nine sources had error covariance matrices that were not positive

definite. This left 10242 sources, of which only 14 involved passive deblending, and these were not used because of the low data count. Thus 10228 matrices were averaged to obtain the results shown in the table below.

	RA	Dec	W1	W2	μ_{RA}	μ_{Dec}
RA	1	-0.0440	-0.0432	-0.0167	-0.8190	0.0369
Dec	-0.0440	1	-0.0031	-0.0068	0.0369	-0.8190
W1	-0.0432	-0.0031	1	0.0036	0.0355	0.0025
W2	-0.0167	-0.0068	0.0036	1	0.0137	-0.0056
μ_{RA}	-0.8190	0.0369	0.0137	0.0098	1	-0.0478
μ_{Dec}	0.0369	-0.8190	0.0025	-0.0056	-0.0478	1

The standard deviations about these means are:

	RA	Dec	W1	W2	μ_{RA}	μ_{Dec}
RA	0	0.0355	0.0791	0.0322	0.0169	0.0305
Dec	0.0355	0	0.0962	0.0381	0.0310	0.0153
W1	0.0791	0.0962	0	0.0140	0.0657	0.0796
W2	0.0322	0.0381	0.0140	0	0.0272	0.0316
μ_{RA}	0.0169	0.0310	0.0657	0.0272	0	0.0382
μ_{Dec}	0.0305	0.0153	0.0796	0.0316	0.0382	0

These statistics show that the proper motion on a given axis is strongly negatively correlated with the position on that axis, but all other proper motion correlations are weak on average and probably not statistically significant.

The most interesting source in this field is the Y dwarf, WISE J0410+1502. This is source number 2651 in the list. The comparison of v6 results to those of v5 are shown below, where the first row in each case is from v6, the second from v5.

```
| v|  src |   ra       |   dec      |sigra   |sigdec  |sigradec| w1snr| w2snr|
 v6   2651  62.5946711  15.0468181   0.2320   0.2454  -0.0438    0.8   25.1
 v5   2651  62.5948188  15.0466549   0.1370   0.1467  -0.0280    1.5   24.8

| v|  w1flux   |  w1sigflux  |  w2flux   |  w2sigflux  |
 v6  2.7288E+00    3.5324E+00  1.6788E+02    6.6758E+00
 v5  5.4330E+00    3.5353E+00  1.6523E+02    6.6743E+00
```

```
| v|   w1rchi2|   w2rchi2|    rchi2 |
 v6  9.060E-01  9.603E-01  9.317E-01
 v5  8.909E-01  1.026E+00  9.573E-01
```

The position uncertainties are a little larger in v6. Opposite effects in W1 and W2 are seen in the photometry, with lower W1 flux and S/N in v6, but higher W2 flux and S/N. In both cases, W1 was not detected on the coadds, as reflected by its low S/N estimates. The flux uncertainties go the other way, with v6 having slightly lower W1 uncertainty and higher W2 uncertainty. The reduced chi-squares follow the S/N pattern, but the overall reduced chi-square is slightly better in v6 despite the inclusion of more parameters in the model. The v6 result for the proper motion is:

```
|  PMRA    |  PMDec   | sigPMRA|sigPMDec|
    1.4398   -1.6355   0.5177   0.5694
```

The units are arcsec/year. The independent measurement of the total proper motion is 2.6 arcsec/year, while the v6 value is about 2.18 arcsec/year; the difference amounts to 0.546-sigma without taking the independent measurement uncertainties into account.

An attempt was made to match UCAC4 sources to the WISE extractions and correlate the proper motions for these associations. 1773 associations were made. No significant correlation was found, but this is probably due to the overwhelming preponderance of sources for which the WISE proper motions are not statistically significant. The reduced chi-square of 4.07 suggests that the WISE proper-motion uncertainties need to be doubled.

Last update - 24 October 2012

John W. Fowler - IPAC (with help from Ken Marsh)