

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Photometry in UV astronomical images of extended sources in crowded field using deblended images in optical visible bands as Bayesian priors

D. Vibert, M. Zamojski, S. Conseil, A. Llebaria, S. Arnouts, et al.

D. Vibert, M. Zamojski, S. Conseil, A. Llebaria, S. Arnouts, B. Milliard, M. Guillaume, "Photometry in UV astronomical images of extended sources in crowded field using deblended images in optical visible bands as Bayesian priors," Proc. SPIE 7246, Computational Imaging VII, 72460U (2 February 2009); doi: 10.1117/12.810228

SPIE.

Event: IS&T/SPIE Electronic Imaging, 2009, San Jose, California, United States

Photometry in UV astronomical images of extended sources in crowded field using deblended images in optical visible bands as bayesian priors

Vibert D.^a, Zamojski M.^b, Conseil S.^a, Llebaria A.^a, Arnouts S.^c, Milliard B.^a, Guillaume M.^d

^a Laboratoire d'Astrophysique de Marseille, OAMP, Université Aix-Marseille & CNRS, Marseille, France;

^b California Institute of Technology, Pasadena, United States

^c Canada France Hawaii Telescope, Kamuela, Hawaii, United States

^d Institut Fresnel, Université Aix Marseille III, Marseille, France

ABSTRACT

Photometry of astrophysical sources, galaxies and stars, in crowded field images, is still a challenging goal, as new space survey missions are launched, releasing new data with increased sensibility, resolution and field of view. The GALEX mission, observes in two UV bands and produces deep sky images of millions of galaxies or stars mixed together. These UV observations are of lower resolution than same field observed in visible bands, and with a very faint signal, at the level of the photon noise for a substantial fraction of objects. Our purpose is to use the better known optical counterparts as prior information in a Bayesian approach to deduce the UV flux.

Photometry of extended sources has been addressed several times using various techniques: background determination via sigma clipping,¹ adaptative-aperture,² point-spread-function photometry^{3,4} isophotal photometry,^{5,6} to lists some. The Bayesian approach of using optical priors for solving the UV photometry has already been applied by our team in a previous work.^{7,8} Here we describe the improvement of using the extended shape inferred by deblending the high resolution optical images and not only the position of the optical sources.

The resulting photometric accuracy has been tested with simulation of crowded UV fields added on top of real UV images. Finally, this helps to converge to smaller and flat residual and increase the faint source detection threshold. It thus gives the opportunity to work on 2nd order effects, like improving the knowledge of the background or point-spread function by iterating on them.

Keywords: UV astronomical imagery, photometry, Data fusion, Bayesian estimation, Point spread function modeling, minimization

1. INTRODUCTION

Launched in 2005 the GALEX instrument⁹ on board of a NASA Explorer mission, is collecting an unprecedented set of astronomical UV images and spectral data in the far (135-175 nm) and the near (170-275 nm) UV range. The telescope continuously scans circular areas of sky of 1°75 of diameter. The key device in the imaging process is the photon counting camera providing the position of each photon in the focal plane. Knowing the telescope movement the process re-situates positions in the field of view, with a precision in resolution of $\sim 5''$. The field is sampled by a grid of 3840×3840 positions or square pixels, each one of 1.5'' wide, subsampling the PSF by at least a factor 3. The image results from the progressive accumulation of counts in each pixel.

The observations are performed in 3 different modes corresponding to different exposure times. Here we are concerned with the 3rd Deep Imaging Survey (DIS) with a long exposure time ($> 30000s$). The DIS fields

Further author information: (Send correspondence to Didier Vibert)

E-mail: didier.vibert@oamp.fr, Telephone: +33 4 95 044163

Address: Laboratoire d'Astrophysique de Marseille, 38 rue Frédéric Joliot Curie, 13388 Marseille cedex 13, France

Computational Imaging VII, edited by Charles A. Bouman, Eric L. Miller, Ilya Pollak, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7246, 72460U · © 2009 SPIE-IS&T · CCC code: 0277-786X/09/\$18 · doi: 10.1117/12.810228

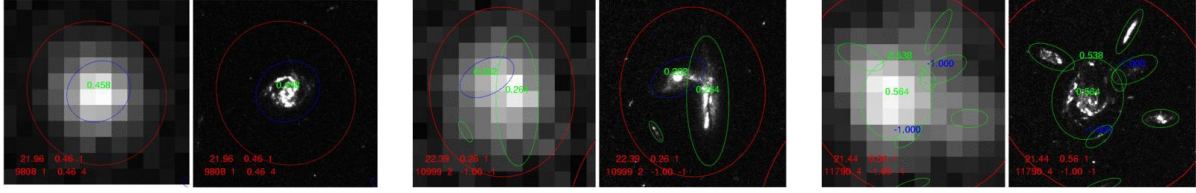


Figure 1. Some examples to compare Galax images (left) and their optical counterpart (right). In these 3 cases, only one object was detected on Galax image using SExtractor where one can see several optical sources that obviously have UV flux.

being very sensitive, contains many faint galaxies below the confusion limit, characterized by blends, ambiguities and miss-identifications of the astronomical sources due to local crowding of very dim galaxies concealed in the background (Figure 1).

Using SExtractor⁶ to extract sources from Galax DIS fields yields to miss a large fraction of objects and moreover to overestimate their flux. The morphological similarity of these UV images to their counterparts in the visible bands suggest to use the visible high resolution counterparts as Bayesian prior information to infer the UV flux. The DAOPHOT software package⁽³⁾ include such an approach but assumes unresolved sources (objects are considered Dirac convolved by the psf) and Gaussian additive noise, whereas Galax images combine extended and unresolved objects and are dominated by Poisson photon shot noise. Then we developed a method to address properly this inverse problem.

In a previous paper⁷⁾ we presented the original procedure used to face on this challenge. This approach uses the Poisson statistics to define the Bayes assumption and solve for the maximum posterior likelihood with the constraint for all fluxes to be positive. Then, in a second paper,⁸ we discussed the photometric performances and behavior of the method when dealing with imperfect knowledge of background, psf and object positions.

Here we will focus on improvements made on the method, using the actual optical image as prior information rather than it's position or elliptical shape obtained from the SExtractor catalog. After a short recall on the EM method we will describe the method to obtain these "stamps", we will also discuss some improvements on psf and re-centering optimization and finally we quantify the gain of doing so.

2. METHOD

2.1 ML parametric Estimation with Priors

The specificity of the proposed photometric procedure is the Bayesian approach under the Poisson noise assumption. The solution is reached with an EM algorithm.

Let be x_i for $i \in \{1, \dots, M\}$ the observed value on pixel i of the UV image considered as a sample of the random variable X_i following a Poisson statistics. Let be $\mu_i = E\{X_i\}$ for $i \in \{1, \dots, M\}$ the expected values for this image. Let be $h_{k,i}$ the known relative value of object $k \in \{1, \dots, K\}$ on pixel i deduced from the visible catalog. Let be $\alpha^T = (\alpha_1, \dots, \alpha_k)$ the vector of unknown fluxes of these objects. Last, let b_i be the known background level value on pixel i of the UV image.

We define the model for the UV image as follows:

$$\forall i \in \{1, \dots, M\} \begin{cases} \mu_i = \sum_{k=1}^K \alpha_k h_{k,i} + b_i \\ P\{X_i = x_i\} = \exp(-\mu_i) \frac{\mu_i^{x_i}}{x_i!} \end{cases} \quad \text{where} \quad h_{k,i} = \sum_j o_{k,j} f_{i-j}$$

here $h_{k,i}$ results from the convolution between each object known profile $o_{k,i}$ with the point spread function f_i of GALEX imaging system. By normalizing the function $h_{k,i}$ to unity, this model allows to estimate directly α , the flux vector of the set of sources in the image.

As was shown in more detail in the first article,⁷ introducing the expectation maximize scheme (EM, see Horiuchi et al., 2000¹⁰) we get the iterative algorithm:

$$\alpha_k^{(n+1)} = \alpha_k^{(n)} \frac{\sum_{i=1}^M (x_i / \mu_i^{(n)}) h_{k,i}}{\sum_{i=1}^M h_{k,i}} \quad \text{where} \quad \mu_i^{(n)} = \sum_{j=1}^K \alpha_j^{(n)} h_{j,i} + b_i \quad (1)$$

The E step compares the data image x_i to the projection $\mu_i^{(n)}$ of the $\alpha_i^{(n)}$ estimates. The result is introduced in the M step as the corrective ratio needed for the new set of $\alpha_k^{(n+1)}$ estimates. To ease the overall algorithm, the background level b is considered known, we found it in a previous procedure, as is described in the article of reference.⁷ The EM application here is closely related to classical Richardson-Lucy methods well know in the literature of image processing for astronomy (see Burud et al,¹¹ Stark and Bijaoui¹²)

The initial guess for α_i is obtained by using the classical pseudo-invert to solve the over determined system:

$$x_i = \sum_{k=1}^K \alpha_k^{(0)} h_{k,i} + b_i \quad \forall i \in x_i > b_i \quad \text{and} \quad \alpha_k^{(0)} \geq |\epsilon(b)| \quad (2)$$

To speed up this step, we can limit ourselves to the diagonal of the pseudo-inverse, which gives:

$$\alpha_k^{(0)} = \frac{\sum_{i=1}^M h_{k,i} (x_i - b_i)}{\sum_{i=1}^M h_{k,i}^2} \quad \forall i \in x_i > b_i \quad \text{and} \quad \alpha_k^{(0)} \geq |\epsilon(b)| \quad (3)$$

In order to tract the large size of the problem (4000 pixels square image and more that 100 000 objects), the entire field is divided in a set of small rectangular “tiles” of identical sizes. To deal with borders, a margin of half the size of the PSF is added around each tile and the margins are overlapping from one tile to a neighboring tile. The priors in the margins are only used to reduce border effects on the objects inside the tile.

2.2 Determining priors information

Before being able to run such an algorithm one need to determine the prior information: basically b_i the background and $h_{k,i}$ the contribution of each objects to each pixels, composed of the PSF f_i and the object known profile $o_{k,i}$.

2.2.1 The objects profile

In the previous implementations, the object profiles could be of two kinds: for a star like object or any unresolved objects, the object was modeled by a Dirac at the exact sub-pixel position given in the optical catalogue, and the psf was interpolated to be shifted at the correct sub-pixel position. For resolved objects, we used the elliptical shape coming out from the optical SExtractor catalogue, rescaled at the Galex resolution and applied an exponential profile along the radius of the ellipses:

$$P = \left(\frac{ab}{\pi/2}\right) \exp\left(-\sqrt{\left[\frac{(x_p - x_c) \cos \theta + (y_p - y_c) \sin \theta}{a}\right]^2 + \left[\frac{-(x_p - x_c) \sin \theta + (y_p - y_c) \cos \theta}{b}\right]^2}\right) \quad (4)$$

where (x_p, y_p) and (x_c, y_c) represent the position of the pixel and that of the object center, and a, b represent the semi-major and semi-minor axis of the object ellipse.

For this framework we described the resolved objects more precisely using the actual optical image rather than the catalog. The idea is simply to degrade the resolution of the image to that of the GALEX image and to convolve each object with the GALEX PSF. However to be able to do this, we must derive the profile of every object independantly meaning that the optical image needs to be segmented and, most importantly, deblended first. Our approach to resolve this problem has been to use a segmentation and deblending method much similar to the one SExtractor uses. Below we describe in more details the method.

2.2.2 Deblending Optical shapes

In order to segment the optical image, we use SExtractor ellipses to define the contours of our objects. We then deblend overlapping objects by dividing the flux of blended pixels among all objects whose contours encompass those pixels in such a way that the fraction of the flux assigned to each object is proportional to the flux that object possesses in the pixel opposite (symmetrically with respect to the center) to the blended pixel. In the event this is not possible because one or more of the symmetric pixels are themselves blended, the objects which do have well-defined fluxes at the location opposite to the blended pixel, are assigned the same flux, at their blended location, as that of their symmetrically opposite pixel. In the case that only one object remains un-deblended at a given location, it is assigned the balance of the flux in that pixel. After this, if there is more than one object that remains to be deblended, then the algorithm proceeds to other pixels, until all pixels have been attempted deblending, at which points it iterates the above process so that newly deblended pixels can be used to further deblend the ones that were not completely successful in the previous iterations. So it goes until either all pixels have been deblended or until no further deblending occurs. In the event that the procedure stops before all pixels have been successfully deblended, the remaining pixels have their flux split among the objects that overlap them in a way that the fraction assigned to each object is proportional to the same exponential of elliptical radius explained in 4. This last step is more arbitrary than the first part of the procedure, but is necessary to bring the deblending process to completion. Typically, in all but the most crowded fields, $\sim 0.1\%$ of the initially blended pixels will have to be deblended this way, and this fraction even often reaches zero.

Figure 2 illustrates the results of this deblending procedure. It shows how objects can be extracted from an optical image and deblended into separate elliptical disks. The original optical image can then be interpreted as a sum of those separate disks, or more succinctly, as the sum of layers in which no two objects overlap.

Once the optical image deblended, we degrade the resolution of each object to that of GALEX, and convolve the degraded image with the GALEX PSF, thus obtaining a GALEX prior profile for each object. These profiles will then be feeded to the EM algorithm for solving the flux. It will determine the weight of each object contributing to the total UV flux in each pixel that has the highest statistical likelihood given the observed data.

2.2.3 The PSF

The precise modeling of the PSF is crucial for the successful issue of the method. The detector structure for the NUV channel sustains a PSF model with a narrow kernel of 3 pixels FWHM surrounded by a very large halo up to 64 pixel of radial extension. The high dynamic range of more than 10^3 in the GALEX images requires a precise determination of this halo in order to minimize the errors due to the halos of brightest stars on the field. The standard modeling using the SExtractor package was revealed inefficient to restore accurately the halo in particular for the far UV channel where the S/N is very low. As alternative we decided to use a two steps approach. In the first step we look for a precise estimation of the kernel using the StarFinder¹³ package. It allows for a precise re-centering and selection. In the second step the PSF is fitted by the sum of three centered Gaussians functions. The dramatic reduction in the degrees of freedom to 3×3 parameters ensure a smoothness and denoising of the PSF.

Despite this effort in the PSF modeling, we know that the PSF determination is not perfect as can be seen in the residual image. Once the objects profiles inferred from high resolution optical bands are included, we see in the residual (the best EM-fitted model subtracted to the UV observed data) a systematic pattern on all bright enough objects. This clearly shows we subtracted a PSF too large. This arises for two reasons mainly. One is the optical PSF (5 times smaller than the GALEX one) that was not deconvolved from the optical image before building the prior shapes. The second cause is due to the stacking of PSFs done by StarFinder. The stacking is performed on manually selected unresolved-objects, whose image is supposed to contain only the PSF. Obviously it is hard to ensure this condition: the object might be slightly extended (a few pixels) as it is the case for many galaxies in the GALEX DIS fields. Then the recentering needed to stack all the PSFs has a limited precision and theses small errors in the PSFs centers, lead to a larger averaged PSF.

In order to correct for this imperfect knowledge of the PSF, we included an optimization step in the algorithm: after a first bunch of EM-iterations computed with the initial estimates of the PSF, we minimize the likelihood with fixed fluxes and varying PSFs parameters. In a first implementation we parameterized the initial psf with only one parameter, doing an homothecy or dilation of the psf. We are looking forward to implement the

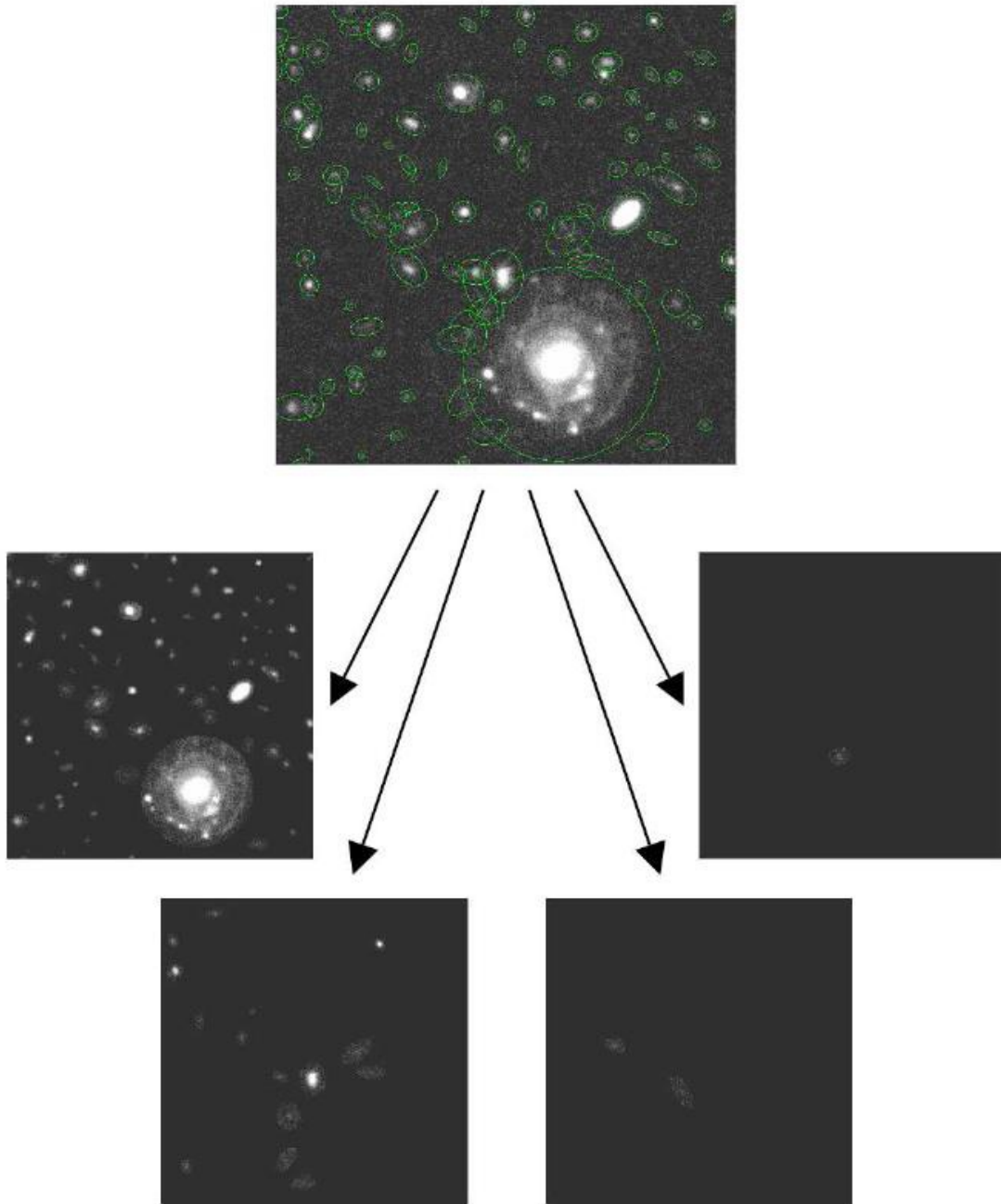


Figure 2. Example of an optical image with SExtractor segmentation ellipses and our resulting debledging shown in four recomposed layers of non-overlapping objects.

optimization on the whole model described above in a future work. Here, for the spatial rescaling of the psf, we employed Brent's method¹⁴ to minimize a function by using parabolic interpolation. This operation is done tile by tile. One need then to gather the PSF estimation on each tile into one PSF non-varying over the field. Once the optimization of the PSF is performed we recompute a bunch of EM-iterations to solve again the fluxes with the newly obtained PSF.

2.2.4 The astrometry

We have seen in the previous works⁸ that the errors on the positions of the priors have big consequences on the precision of the resulting photometry. We thus tried to improve the correct centering of priors in the field.

The first thing to achieve before any further processing is a correction of the distortion of the Galex image. To do so we cross-correlate the positions of the brightest objects detected with SExtractor in the Galex image with the brightest objects of the prior catalogue. Supposing the positions of the prior catalogue have negligible errors (recall they are obtained from a lot higher resolution images) we then fit the discrepancies of the two catalogues by 2^{nd} degree polynome on x and y . Then all the positions of the prior objects are warped using this polynome.

The second correction for astrometry is then performed at the same stage where we optimize for the PSF. After a first bunch of EM-iterations, we use the estimated image to correct the recentering. Either we compute a sub-pixel (dx, dy) displacement on each tile via oversampling and FFT cross-corelation of the observed image and estimated one. Or we minimize the likelihood with respect to a (dx, dy) displacement of all the priors.

2.2.5 The background

Several choices can be made here. The first one is to use the background determined by the Galex data reduction pipeline. It has the advantage of being free from non-linear filtering of the original image but with the drawback of being very smooth.

The other is to use the background as determined in⁸ using the inferior hull of the image found through a nonlinear multiresolution top-down approach. It has the advantage of containing more high-frequency structures with the drawback of possibly filtering out the unknown signal we are trying to recover.

The third is to use the low-frequency background knowing, it may lack some high frequency components, and try to recover these sharp components from the residual after applying the EM-iterations. Note that we are talking here of frequency components that need be larger than the noise frequency components. The full process in then:

1. run EM solving with an initial background
2. mask-out from the residual a small region (smaller than the PSF) around all the objects (since a small error in flux or prior position leads to an artifact in the residual).
3. fill the masked region using the multiresolution method cited above.
4. low-pass filter out the Poisson noise high frequency components
5. redo the EM-iteration with the new background

2.2.6 The error estimation

There is no known analytical expression for the variance of the estimated flux that minimize the likelihood. We thus empirically quantify the variance $\hat{\sigma}_k$ for the flux of object k by the rms, under the objects area, of the residual (best fitted model subtracted to the observed image) weighted by the contribution of the considered object in every pixel and normalized by the sum of the square of the contributions:

$$\hat{\sigma}_k^2 = \frac{\sum_{i=1}^M h_{k,i}(x_i - \hat{\mu}_i)^2}{\sum_{i=1}^M h_{k,i}^2} \quad (5)$$

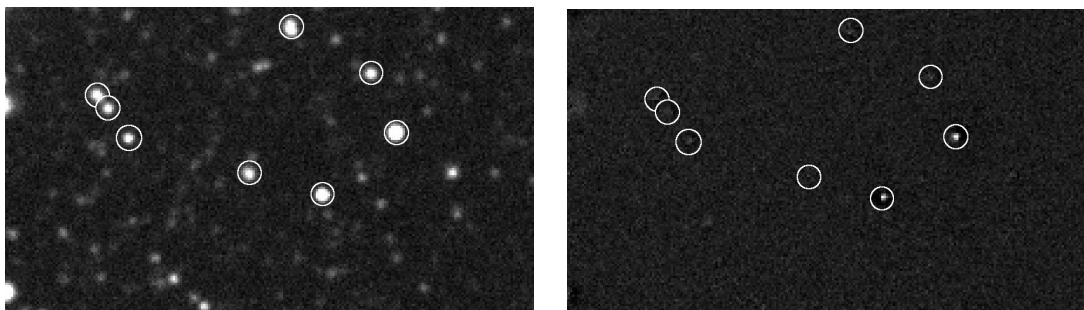


Figure 3. Estimation of the variance on the estimated fluxes (*left*) from the weighted mean square of the residual (*right*).

where $\hat{\mu}_i = \sum_{j=1}^K \hat{\alpha}_j h_{j,i} + b_i$ is the estimated value on pixel i for the estimated fluxes $\hat{\alpha}_j$. We show that this estimation is consistent with the true errors obtained from the simulation. The figure 5 shows how the quality of the residual under objects area may be used to estimate the variance.

3. RESULTS

The figure 4 shows an example of the method applied to the Galex XMMLSS Near UV field using the CFHTLS data as optical priors. The residual shown is a lot better compared to what could be obtained before using the full optical image as priors and optimizing the PSF.

The behavior of the method is quantified using simulations. We inserted randomly in the field a large set of pseudo-stars on the original UV images. These star-like objects were built with the PSF deduced from the original UV field itself. This method allows us to determine the error of magnitude vs. magnitude relationship as well as the operational range of magnitudes. The figure 5 shows the recovered fluxes by bins of magnitude. It shows that up to magnitude ~ 25.5 which is one magnitude higher than the original Galex catalogue the method is complete. There is still an Eddington bias in the recovered distribution of magnitude, inherent to the asymmetry of the input distribution and hard to eliminate totally.

The estimated variance from the residual on these known simulated objects appears fully consistent with the dispersion shown in figure 5.

4. CONCLUSION

Test results show a noticeable gain on magnitude limit as well as in resolution relative to the standard pipeline method using SExtractor. On current DIS fields, EM provides good photometry and completeness down to magnitude 25.5 which is 1 magnitude deeper than the GALEX pipeline.

The use of shaped priors combined with a simple psf optimization improved drastically the residual. If less impressive, the photometry has been improved also. The dispersion has been reduced for the faint magnitudes (~ 25).

The gain on the residual will now let us work on the background by iterating on it after masking and filtering the residual. By doing so, we may gain sharp features that may have interesting astrophysical sense.

REFERENCES

- [1] Costa, G. D., “Basic photometry techniques,” in [*Astronomical CCD Observing and Reduction Techniques*], Howell, S. B., ed., *ASP Conf Series* **23** (1992).
- [2] Kron, R. G., “Photometry of a complete sample of faint galaxies,” *The Astroph. Journal* **43**, 305–325 (June 1980).
- [3] Stetson, P., “Daophot a computer program for crowded field stellar photometry,” *Publ. Astron. Soc. of the Pacific* **99**, 191 (1987).

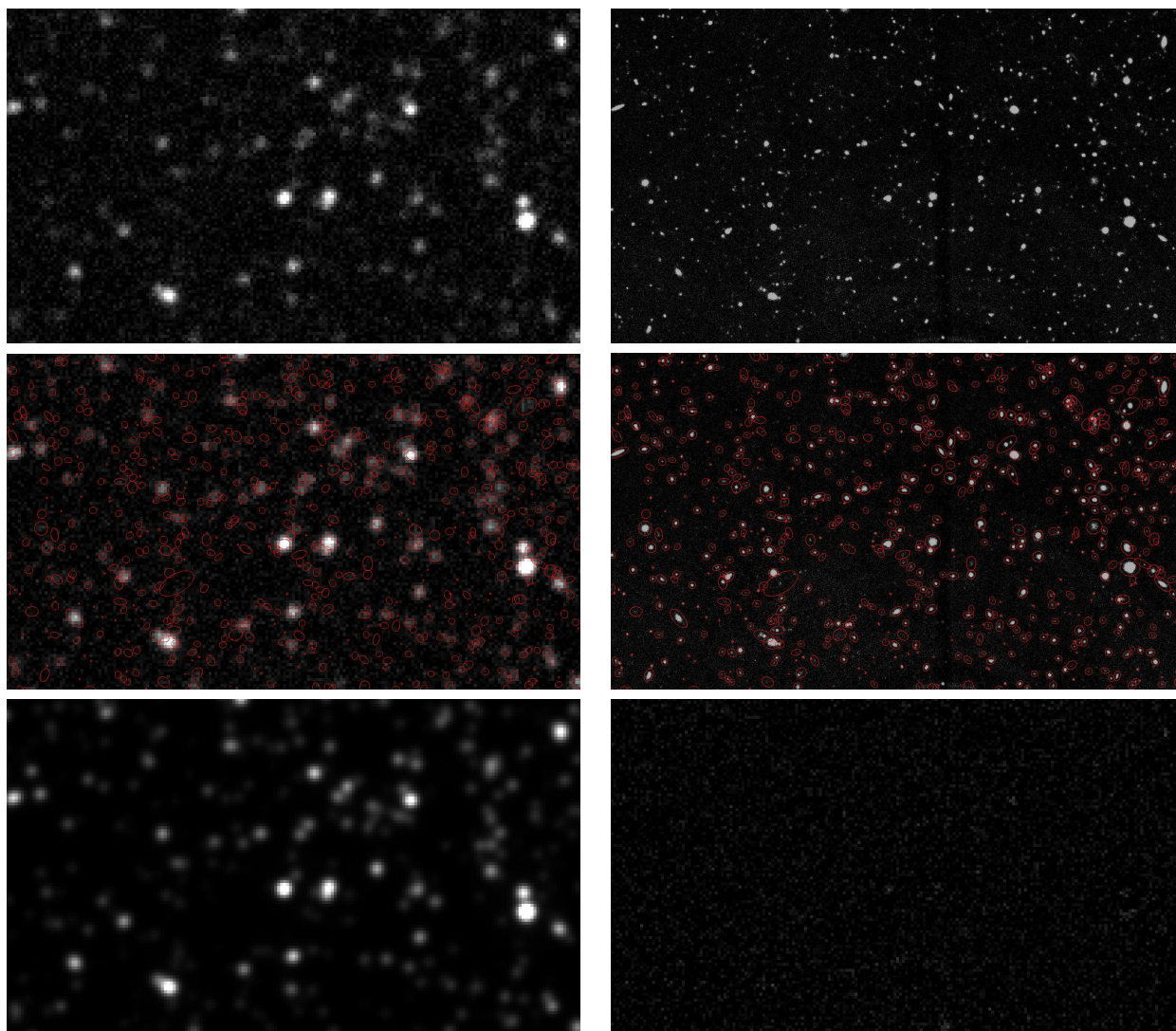


Figure 4. Example of 100 EM iteration using deblended optical shapes. *Top left*: GaleX XMMLSS field. *Top right*: corresponding CFHTLS optical U band. *Middle left*: same GaleX field with superposed selected priors. *Middle right*: CFHTLS selected sources with magnitude < 26.5 . *Bottom left*: optimal solution from EM convergence. *Bottom right* residual

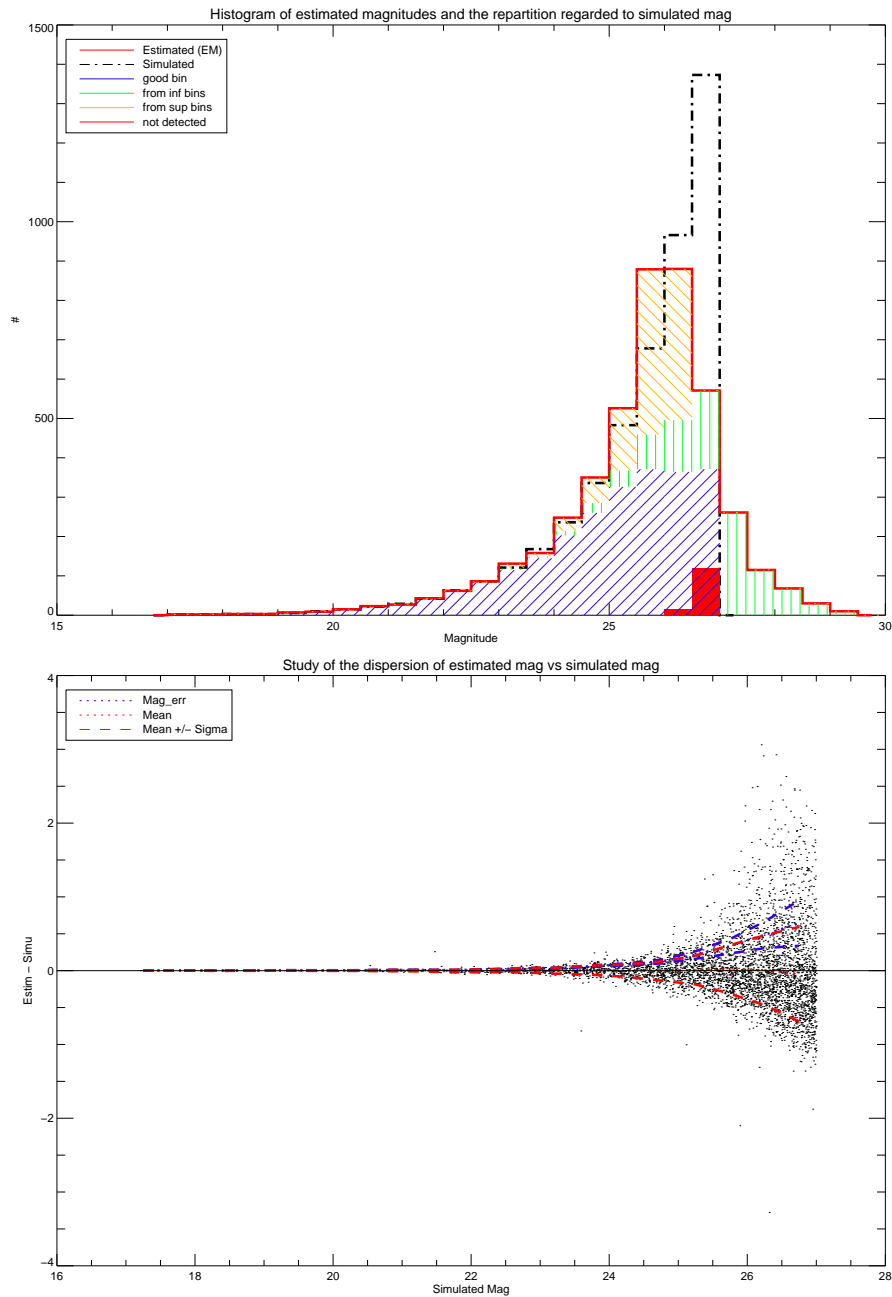


Figure 5. Quantitative plots showing the recovered fluxes, expressed in magnitudes, of randomly added simulated sources on a true Galex image. *Top*: dashed line is the input distribution and plain line is the recovered one. The small plain filled portion of the histogram shows the undetected objects. The bottom diagonally filled portion shows the sources detected in the correct bin, whereas the vertically filled portion are the sources overestimated and the top diagonally filled portion are the underestimated ones. *Bottom*: scatter plot of the recovered fluxes with the estimated variance is exactly superposed on the 1σ dispersion.

- [4] Penny, A. J. and Dickens, R. J., “CCD photometry of the globular cluster NGC 6752,” *MNRAS* **220**, 845–867 (June 1986).
- [5] Beard, S. M., MacGillivray, H. T., and Thanisch, P. F., “The Cosmos System for Crowded-Field Analysis of Digitized Photographic Plate Scans,” *MNRAS* **247**, 311–+ (Nov. 1990).
- [6] Bertin, E. and Arnouts, S., “SExtractor: software for source extraction,” *Astron. Astrophys. Suppl. Series* **117**, 393–404 (1996).
- [7] Guillaume, M., Llebaria, A., Aymeric, D., Arnouts, S., and Milliard, B., “Deblending of the uv photometry in galex deep surveys using optical priors in the visible wavelengths,” in [*Image Processing: Algorithms and Systems, Neural Networks and Machine Learning*], Dougherty, Astola, Egiazarian, Nasrabadi, and Rizvi, eds., *Proc. SPIE* **6064**, 11–1–11–10 (2005).
- [8] Llebaria, A., Magnelli, B., Arnouts, S., Pollo, A., Milliard, B., and Guillaume, M., “Multi-channel 2D photometry with super-resolution in far UV astronomical images using priors in visible bands,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **6812** (Mar. 2008).
- [9] Martin, D., Fanson, J., Schiminovich, D., Morrisey, P., Friedman, P., Barlow, T., Conrow, T., R. Grange, Jelinsky, P. N., B. Milliard, Siegmund, O., Bianchi, L., Byun, Y., J. Donas, K. Forster, M. Heckman, T., Lee, Y.-W., Madore, B., Neff, S., Rich, R., T. Small, F. Surber, A. Szalay, B. Welsh, and Wyder, T., “The galaxy evolution explorer: A space ultraviolet survey mission,” *The Astroph. Journal Letters* **619**, L1–L6 (2005).
- [10] Horiuchi, S., Kameno, S., and Ohishi, M., “Developing a wavelet clean algorithm for radio-interferometer imaging,” in [*Astronomical data analysis software and systems X*], F.R. Handen, F. P. and Payne, H., eds., *Conference Series* **238**, 529–532, The Astron. Soc. of the Pacific (2001).
- [11] Burud, I., Stabell, R., Magain, P., Courbin, F., Ostensen, R., S. Refsdal, and Remy, J., “Three photometric methodes tested on bround based data of q2237+0305,” *Astron. Astrophys.* **339**, 701–708 (1998).
- [12] Stark, J. and Bijaoui, A., “Multiresolution deconvolution,” *J. Optical Soc. Am. A* **11**(5), 1580–1588 (1994).
- [13] Diolaiti, E., O. Bendinelli, Bonaccini, D., Close, L., Currie, D., and Parmeggiani, G., “Starfinder,” in [*Adaptive Optical Systems technology*], Wizinowich, P., ed., *Proc. SPIE* **4007**, 879–888 (2000).
- [14] Brent, R. P., [*Algorithms for Minimization without Derivatives*], Prentice-Hall (1973).