X-ray Source Detection in EPIC Images using Wavelet Transforms

F.Damiani, A.Maggio, G.Micela and S.Sciortino
Osservatorio Astronomico di Palermo G.S.Vaiana, Piazza del Parlamento 1, Palermo, ITALY

ABSTRACT

We present applications to simulated XMM/EPIC images of a source detection code, based on wavelet transforms, originally developed for the analysis of ROSAT PSPC/HRI images. Because of the much higher sensitivity anticipated for the XMM EPIC camera with respect to the ROSAT PSPC, one expects to detect a much larger number of sources per square degree, most of which near or just above the detection threshold. EPIC images of moderate exposure time will therefore reach rather easily the confusion limit. Both these difficulties (source confusion and presence of many weak sources) are efficiently handled using the wavelet transforms, that have proven to be effective both in detecting weak sources and in resolving close groups of point sources. Also extended sources are efficiently detected with this method, and recognized as such. We expect to detect a large number of weak sources in the Pleiades and IC 2391 open clusters and in the Sco-Cen association as well as in the many other young clusters, that will be observed as part of the XMM GTO program. Detection of weak (low-mass) emitters is crucial to determine the initial mass funtion at low-mass end.

With this detection algorithm, we expect to be able to take full advantage of the information content of EPIC images (as well as of the AXAF ones, for which a similar detection algorithm is currently being developed by our team). This will allow e.g. to study the X-ray flux distribution of sources down to a very low flux limit (the ability of reconstructing an input log N-log S from the output of the wavelet detection algorithm has been already studied by Damiani et al. (1998) for the PSPC case), enabling a much more accurate and representative census of samples of both galactic and extragalactic object populations to be obtained.

1. Outline

Wavelet Transforms (WT) are a proven efficient tool to select interesting features in astronomical images. Their most basic application in this context is perhaps the detection of sources in an image, in which they excel also since they allow a multiscale analysis of the data. This is of great advantage in the detection of extended sources, or of point sources broadened by the instrumental PSF. X-ray source detection methods based on WT or similar techniques have been presented by Rosati et al. (1993), Grebenev et al. (1995), Vikhlinin et al. (1994), and Damiani et al. (1997a,b).

The WT of an image is a function $w(x, y, a)$ depending on both position (x, y) and a scale parameter a, and is defined as:

$$w(x,y,a) = \iint g\left(\frac{x-x_0}{a}, \frac{y-y_0}{a}\right) f(x_0,y_0)\, dx_0 dy_0 \tag{1}$$

where $g\left(\frac{x}{a}, \frac{y}{a}\right)$ is the 'generating wavelet', for which we choose the so-called 'Mexican Hat" function:

$$g\left(\frac{r}{a}\right) = \left(2 - \frac{r^2}{a^2}\right)\, e^{-\frac{r^2}{2a^2}} \qquad (r^2 = x^2 + y^2). \tag{2}$$

Among the properties of the WT, we remind those most relevant for source detection: The WT of a uniform background is zero (the same is true for a uniform background gradient). The function $w(x, y, a)$ depends on values of the analyzed function f (the data) only within 4-5a from the point (x, y). The WT at a particular scale a brings into evidence image structures of size $\sim a$. Moreover, the WT analysis allows to deblend partially overlapping sources, if they are intense enough.

As detailed in Damiani et al. (1997a), our approach takes properly into account the transformation of the Poisson photon statistics of the original image into a different statistics in the wavelet-transformed image. These statistical properties depend only on the functional form of the generating wavelet, and not on the specific dataset under examination. This allows us to set accurate thresholds for detection at each wavelet scale.

Our algorithm runs through the following steps:

1) Determination of local, scale-dependent background.

2) Putative source recognition in the WT images and assessment of source existence based on local background.

3) Determination of source size and intensity in the WT space.

4) Iteration of steps 1-3, with an improved source-free background map.

2. Implementation of the Algorithm

- **Background Map:** The reference background used to predict the expected WT fluctuations in the absence of sources is computed by smoothing the image using a gaussian filter. In each point, the width of the gaussian has been taken equal to the local PSF width σ_{PSF}: this ensures that the resulting map is a faithful representation of the image background, describing *real* background modulations on any length scale, and least affected by statistical errors.

Since this map includes the effect of sources, a further step is to apply a local median filter that minimizes their effect on the final reference background. The region used to compute medians has a size proportional to the WT scale a, in order to have a fully scale-invariant approach.

- **WT Computation:** The WT of the image is computed at each scale a (by direct integration), over a region of radius such that $a \geq \sigma_{PSF}$ (since no information is present in the data on scales $a \ll \sigma_{PSF}$), and with a grid spacing proportional to a, not to oversample the WT image.

- **Source Detection:** Candidate sources are identified as local *spatial maxima* of the WT at each scale. A candidate source is retained as a real detection if its WT amplitude w_{peak} is higher than that expected for the WT of the local background (at that scale a), to an assigned probability level.

- **Source Count-Rate and Size Determination:** Sources detected at various scales are cross-identified, to yield the profile of WT amplitude as a function of scale a, for each source. This is used to derive the *source rate* and *size*. At the scale of maximum significance, the WT is recomputed in a small region around the peak, using a finer spatial grid, to best estimate the source position and count rate. Errors on count rate and size are computed also from the error in the value of $w(a)$, but in doing this we must treat appropriately datapoints that are not statistically independent (for details see Damiani et al. 1997a).

- **Background Map Update:** After having obtained a first list of sources, a second iteration consists in the elimination of all detected point sources from the computation of the reference background map; this step allows us to detect even weak sources close to much stronger ones, which spuriously raise the local background estimate. Having updated the background map, all previous steps are repeated to build a final source list.

3. Upper Limits and Sensitivity Map

The upper limit for an undetected source is computed in nearly the same way as the count rate for a detected source. The upper limit (at a given significance level) depends on the local background and on the source apparent size (and shape). The source size is nonzero even for point sources, in which case it is determined by the local PSF width. In this way, limiting count rates for undetected objects are entirely consistent with the count rates of actual detections.

By computing upper limit values on a closely spaced grid over the field of view, we construct a sensitivity map for the given observation. From this and the associated sensitivity histogram we can then derive the $\log N$-$\log S$ distribution of the detected source population. However, to do this accurately a number of different corrections must be applied, as explained in Damiani et al. (1998).

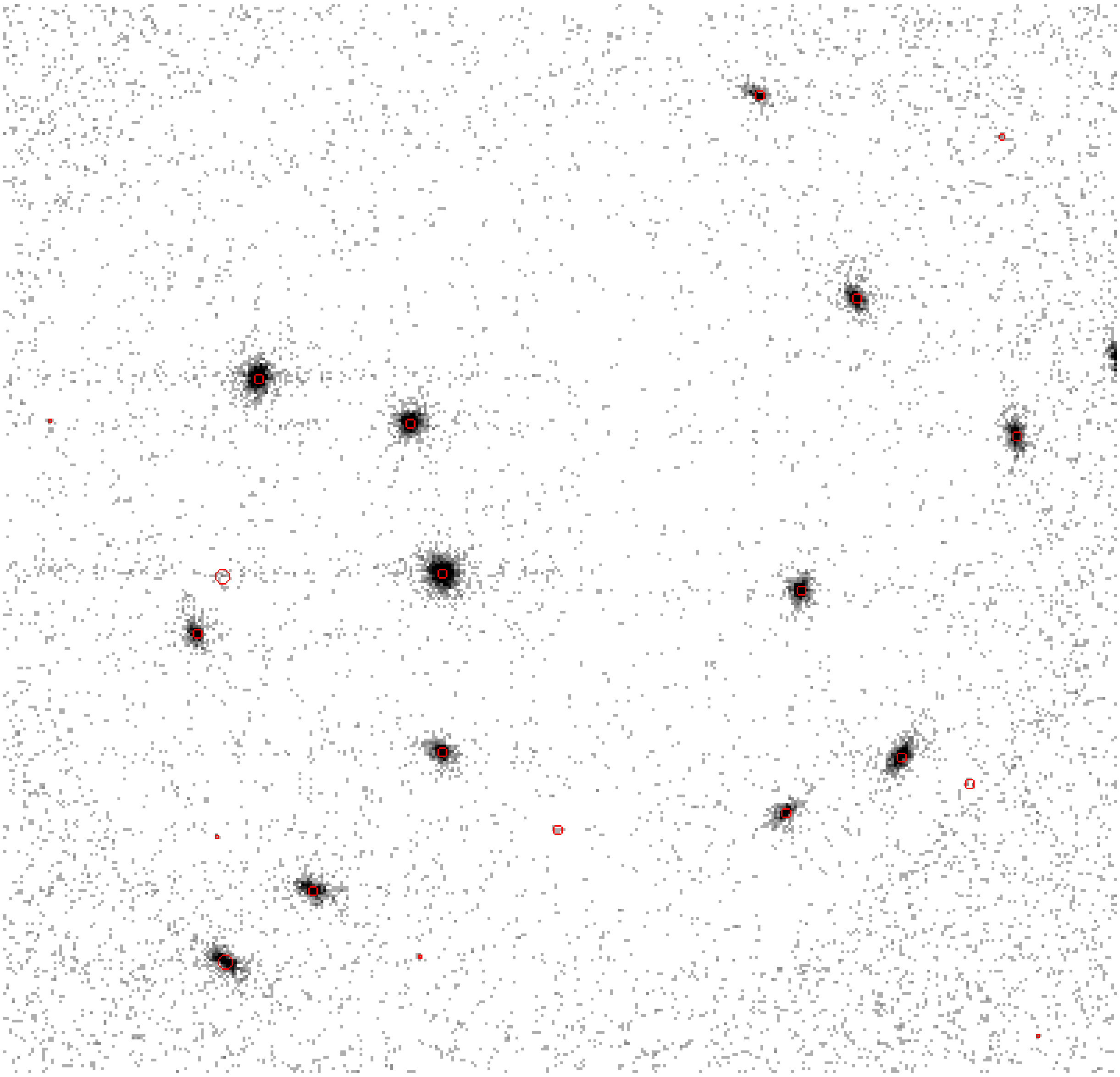

Fig. 1.— EPIC-PN simulated image of part of the Pleiades open cluster. The exposure time is 30 ksec. Red circles indicate wavelet-detected sources, the radius of each circle being equal to the detected source size.

4. Testing on Simulated EPIC-PN Images

In Figures 1 and 2 we show two 30-ksec simulated EPIC-PN images of the Pleiades cluster, and of the σ Ori cluster, obtained using the SciSim simulator. Figure 3 is an enlargement of a partially confused region of Figure 2. Sources detected with our algorithm are shown as circles, whose radius is proportional to the detected source size. The scales used range from 5 to 16 arcsecs, and would allow therefore to detect and recognize also extended sources (not present in these simulations). The threshold chosen should correspond to the detection *on the average* of one spurious source per field, but detailed simulations to better quantify this threshold are in progress.

In both Figures 1 and 2 long 'tails' attached to the images of bright sources are visible, caused by out-of-time events recorded in the CCD. Such features, for sufficiently bright sources, may lead to the detection of spurious sources all along them (a few of which are visible in the Figures). Ad hoc techniques have to be considered to recognize and avoid these spurious sources.

While Figure 1 shows mostly bright sources, a number of weak sources is instead present in Figure 2 (Orion), of which almost all are detected by our algorithm (even in its current, not finally optimized version). It can also be seen that bright sources tend to be detected at larger scales than weaker sources. This does not happen in simulations of purely gaussian sources, and it must therefore be related to the large scattering wings of the XMM PSF. This is an area in which further improvements in the code are planned. Anyway, as Figure 3 shows clearly, we are able to detect separately all sources in a close group, even though their scattering wings partially overlap.

REFERENCES

Damiani,F., Maggio,A., Micela,G., Sciortino,S. 1997, ApJ, 483, 350.

Damiani,F., Maggio,A., Micela,G., Sciortino,S. 1997, ApJ, 483, 370.

Damiani,F., Micela,G., Sciortino,S. 1998, Astron. Nachr., 319, 78

Grebenev,S.A., Forman,W., Jones,C., Murray,S. 1995, ApJ, 445, 607.

Rosati,P., Burg,R., Giacconi,R. 1994, in *The Soft X-ray Cosmos*, eds. E.M.Schlegel and R.Petre (AIP Conf. Proc. 313), p.260.

Vikhlinin,A., Forman,W., Jones,C., Murray,S. 1995, ApJ, 451, 542.

This preprint was prepared with the AAS LaTeX macros v4.0.

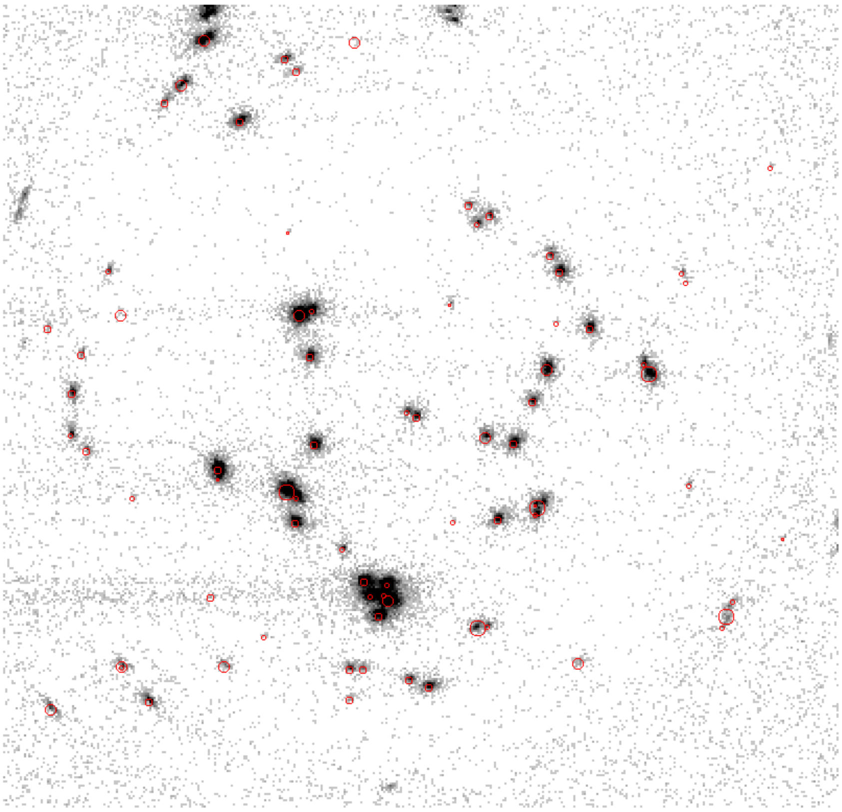

Fig. 2.— EPIC-PN simulated image of the region around σ Ori. The exposure time is 30 ksec. Red circles indicate wavelet-detected sources, the radius of each circle being equal to the detected source size.

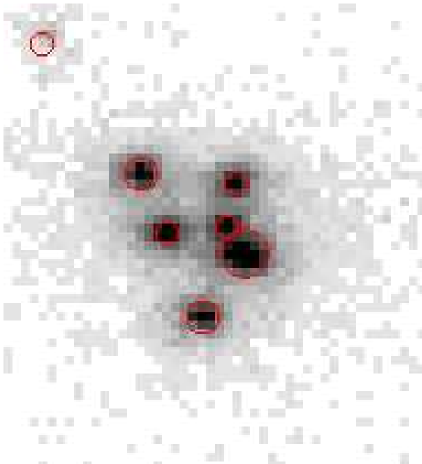

Fig. 3.— Enlargement of a subregion of Figure 2, showing how a close group of sources is resolved by our detection algorithm, despite the partial overlap of the wings of the sources' PSFs.