False Detection Rate of Source Finding Revisited

W. D. Cotton (NRAO) May 24, 2016

***Abstract*—Astronomical images always contain a randomly distributed component to pixel values which is unrelated to anything on the celestial sphere. The statistics of this distribution must be taken into account in order to establish the significance of features in the image. The true distribution of the "noise" can be estimated from the pixels with negative values and used to estimate the probability that a given positive pixel value was obtained at random. "Sources" can be selected which have a probability of a false detection below a selected level. A previous attempt to apply this technique to low frequency radio images produced unsatisfactory results due to the highly variable nature of the images. This memo explores the use of an estimation of the false detection rate for sources as a function of flux density using a well behaved, deep VLA image. Use of a 5% target false detection rate (FDR) produced a sample with an estimated FDR of 7.7% with a 74% increase in the number of sources over a 5 σ selection. Four σ selection gives an estimated FDR of 0.5% with an increase of 38% in the number of sources over a 5 σ sample.**

***Index Terms*—Radio Interferometry, Source cataloging**

I. Introduction

WIDEFIELD astronomical images, especially those produced by sky surveys, are frequently decomposed into a list of discrete components to produce a catalog of objects. Due to a random component of the pixel values not of celestial origin, some criterion must be adopted to distinguish features in the image which correspond to plausibly real objects from those unlikely to be real. In the following this non–celestial component of the pixel value distribution will be called "noise". Such tests must invariably involve a trade off between the possibility of missing real objects and contamination of the catalog by fictitious ones.

Some measure of the statistical probability of a feature being due the the noise distribution is required. Such tests are well established in cases where the noise has a Gaussian distribution. The simplest such test is a cutoff at some multiple of the Gaussian σ of the distribution where values in excess of this are expected to be sufficiently rare. For more sophisticated applications of Gaussian statistics to astronomical images, see [1], [2].

An attempt to select sources based on an estimation of the false detection rate on a low frequency radio sky survey was described in [3]. The results of this attempt were unsatisfactory as the pixel statistics of the low frequency images were sufficiently variable to limit the region over which the statistics were constant to be too small to measure good statistics. The following is a reworking of [3] using well behaved images from a deep single pointing survey with the VLA at S band (2-4 GHz).

W. Cotton is with National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA, 22903 USA email: bcotton@nrao.edu

In some applications, the noise distribution is not well modeled by a Gaussian and Gaussian statistics underestimate the number of false detections. Calibration and imaging artifacts will contribute a non–Gaussian component to the distribution of pixel values and frequently do so by increasing the tails of the distribution.

The following develops a technique of estimating the false detection probability directly from images statistics and applies it to deep VLA images. Here, false detection rate (FDR) is the likelihood that a feature could be a random occurrence due to the noise distribution. The faint sources in the image tested are dominated by star forming galaxies, the majority of which are detectable in deep IR images. Mark Lacy has kindly provided deep Spitzer 3.6 and 4.5 micron images covering the region surveyed as well as a catalog of objects derived from them [4]. A catalog derived from the radio image can be compared with the Spitzer IR catalog to derive a statistical measure of the false detection rate actually achieved. The techniques discussed are implemented in the Obit package ([5], http://www.cv.nrao.edu/∼bcotton/Obit.html).

II. Interferometric Image Statistics

The "noise" in radio interferometer images arises from a number of causes, thermal noise from the atmosphere, antenna, ground pick and electronics as well as power from celestial sources scattered due to incomplete calibration, especially phase calibration. All but the last of these will contribute equally to all portions of the derived image. Other imaging artifacts such as result from limited dynamic range images including bright sources can also contribute to the distribution of pixel values; these artifacts will generally be more concentrated around the bright sources. The spatial filtering inherent in interferometric images means that the noise in the image will be centered near zero.

A. Example Image

For the purposes of the tests described in this memo, data from a deep, single pointing VLA survey at 2-4 GHz are used. Images at 3" and 0.7" FWHM were obtained for the same region of the sky. The 3" images were derived from data in a number of VLA configurations over a number of years while the 0.7" data were taken in a single VLA configuration ("A") over a period of several months. Since the 3" image is being used to derive the catalog of radio sources, most of the following uses it. The uv–coverage used for both these images is excellent which coupled with the absence of strong sources results in images which are not dynamic range limited and have relatively uniform noise properties.

The pixel distributions from these images are given in Figure 1. The positive tail in excess of the negative tail of

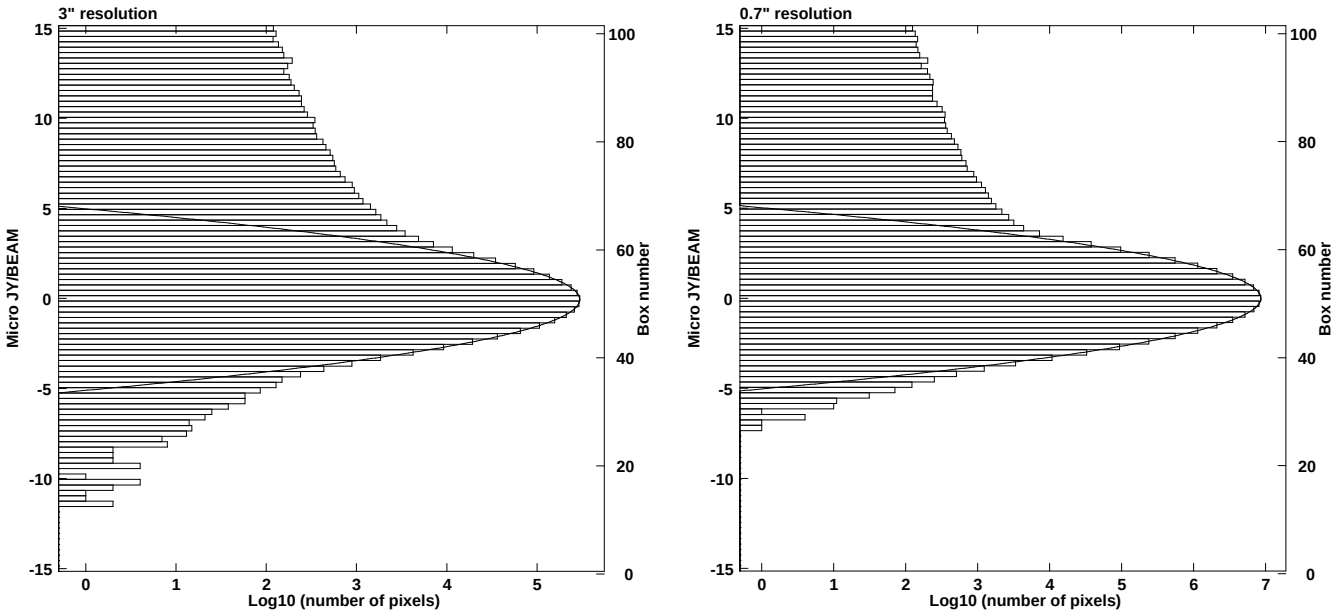

Fig. 1. Pixel histograms for 3" (left) RMS=1.00 μJy and 0.7" (right) RMS=0.89 μJy. Solid lines are the Gaussian fitted near the peak of the distribution. The negative tail of the distributions indicate the non-Gaussian nature of the distribution of values.

the distributions is dominated by real sources; however, the negative tail is expected to include a combination of thermal noise and calibration and imaging artifacts. The negative tail of the distribution from the 3" image (left) shows a distinct deviation from Gaussian, perhaps unsurprising given the period of time and number of observations involved. The distribution from the 0.7" image is better behaved but still deviates from Gaussian well away from the peak (although at a level well below the positive side of the curve. Clearly, Gaussian statistics are not a good model for the distribution of values well away from zero, even in these well behaved cases.

III. Estimation of Actual Noise Distribution

Real sources undoubtedly contribute much of the positive wing of the pixel distributions seen in Figure 1 but the negative wing of the distribution should be an indication of the true noise distribution. If the true positive wing of the noise is symmetric with the negative wing, the negative half of the distribution can be used to estimate the corresponding positive half. Thus, the excess positive values in a given interval over the negative values in the corresponding negative interval gives the fraction of the positive values related to real sources. In this simple case, the false detection rate for positive values in this interval can be expressed as:

$$FDR_x \;=\; 1 \;-\; \frac{n_+ - n_-}{n_-}$$

where FDR_x is the false detection rate at flux density level x, n_+ is the number of pixels in the positive x bin and n_- is the number of pixels in the negative x bin.

Determining flux density levels for low false detection rates requires good statistics well out in the wings of the distribution. This generally means sampling large numbers of pixels and including large areas. On the other hand, the character of the noise may change across the image meaning statistics over more limited areas are preferred. One compromise to make the statistics more robust is to integrate the pixel histogram; each bin then includes the counts in that bin plus all bins further from zero. This makes a subtle change in the meaning of the false detection rate given above; it is the probability of a pixel at a given level, or greater, being a random event. The difference between differential and integrated histograms is illustrated in Figure 2. The flux density corresponding to a given false detection rate can be interpolated from the values shown as "+"es in Figure 2 right.

IV. Implementation

A histogram analysis of the false detection rate as described in the previous section was implemented in the Obit package. In the c library, the ObitPixHisto class implements the functionality with bindings in python as the PixHistFDR class. The source finding task, FndSou was modified to allow selecting sources by estimated false detection rate with statistics derived from a box of a given size centered on the source in question.

A. Cataloging

The 3" image was used by Obit task FndSou to generate catalogs. This program does fitting to islands in the image of at least 2×2 pixels in excess of a given threshold. Elliptical Gaussians are then fitted to the islands. The significance of each fitted component was based on the peak value by two statistical criteria. The first of these criteria is that the peak value exceed some multiple of the RMS in the pixels within a 101×101 pixel box centered on the component. The other criterion uses an estimate of the false detection rate as

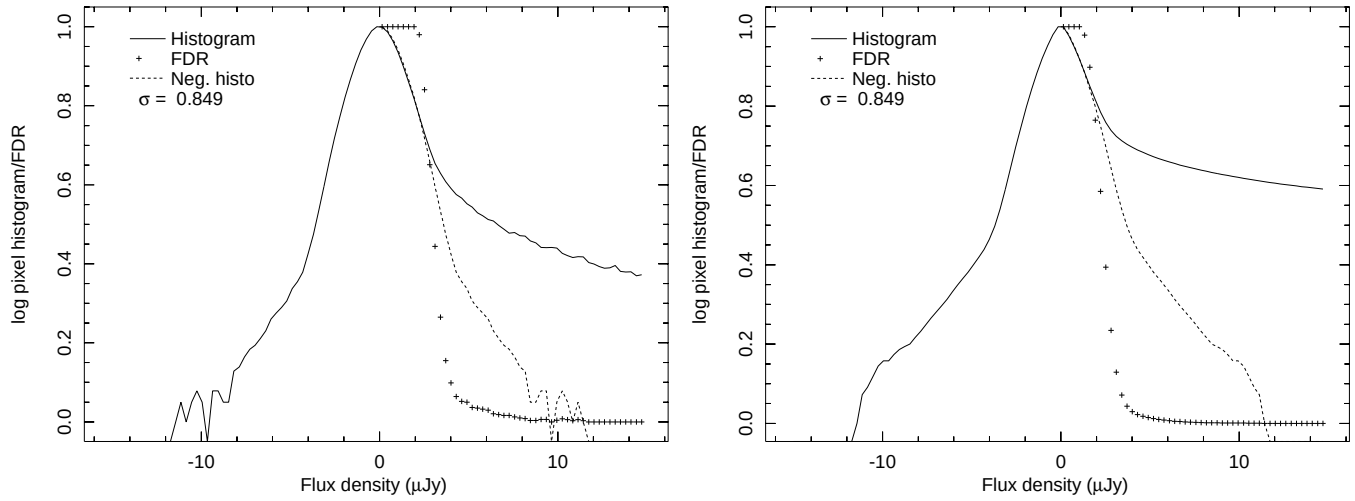

Fig. 2. Differential and integrated histograms from the same set of pixels from the 3" image. The solid line is the histogram normalized to 1, the dashed line is the negative half folded to the positive half and the "+"es are FDR. On the left is the differential histogram and on the right the integrated version.

described above for a given target false detection rate (5%) using 2000×2000 pixel boxes.

B. Comparison of Radio and IR Catalogs

The validity of entries in the radio catalogs were determined by the presence of a match with an entry in the IR catalog. The resolution of the IR images is approximately 2" so the matching was done using the radio 3" image catalog which should also be relatively complete. This test excluded the region around a bright star in the IR images which suffered from saturation and imaging artifacts as well as regions around extended emission in two radio sources. While the vast majority do, not all of the radio sources have detected IR counterparts. The fraction expected to be detected was determined using a high significance sample ($> 7\ \sigma$) assuming that this value does not change over the range of radio flux densities between the fainter and strong samples. The average position offset of the matches in the $>\ 7\ \sigma$ sample was removed when searching for matches.

C. False Matches

The density of IR sources is sufficiently high that there is a finite probability of an IR source within a given distance of a random position. The distribution of nearest IR source to entries in the deepest radio catalog is given in Figure 3 left; the number of matches drops to the base level beyond 2". The distribution of nearest IR sources to an arbitrary grid of positions is given in Figure 3 right. The integrated probability of a random match within 2" from this figure is 14.1%; large enough to affect the statistics of matches. The expected number of non matches is

$$n\ =\ (F + g') \times (1 - f) \times s$$

where F is the achieved false detection rate, g' is the fraction of radio source without a detectable IR counterpart, f is the probability of a false match and s is the total number of samples searched for a match. Solving for F:

$$F\ =\ \frac{n}{(1 - f) \times s} - g'$$

f can be estimated from the histogram in 3 right for a given allowed match distance and g' can be estimated from a strong sample of radio source whose reality is certain but which lack a detected IR counterpart. The estimate of g' will also be affected by false matches as given by:

$$g'\ =\ \frac{g}{1 - f},$$

where g is the ratio of nonmatches to samples compared for a sample of (real) radio sources. The achieved FDR is then

$$F\ =\ \frac{n}{(1 - f) \times s} - \frac{g}{1 - f}. \qquad (1)$$

D. Achieved FDR

Test matches were performed with a number of radio samples selected by a target FDR and an additional SNR test; these are shown in Table I. The column labeled "total" is the total number of entries in the radio catalog including those in the exclusion regions; column "fract" is the fraction of entries in the radio catalog with no match in the IR catalog. Column "FDR" gives the achieved false detection rate as given by Eq. 1. An SNR$>$ 3 limit essentially adds no constraint over the FDR selection. Likewise, the FDR selection had no affect on the SNR$>$ 7 sample. The alignment of a sample of radio sources selected above a FDR of 5% (SNR$>$ 3) with the 4.5 micron image is given in Figure 4.

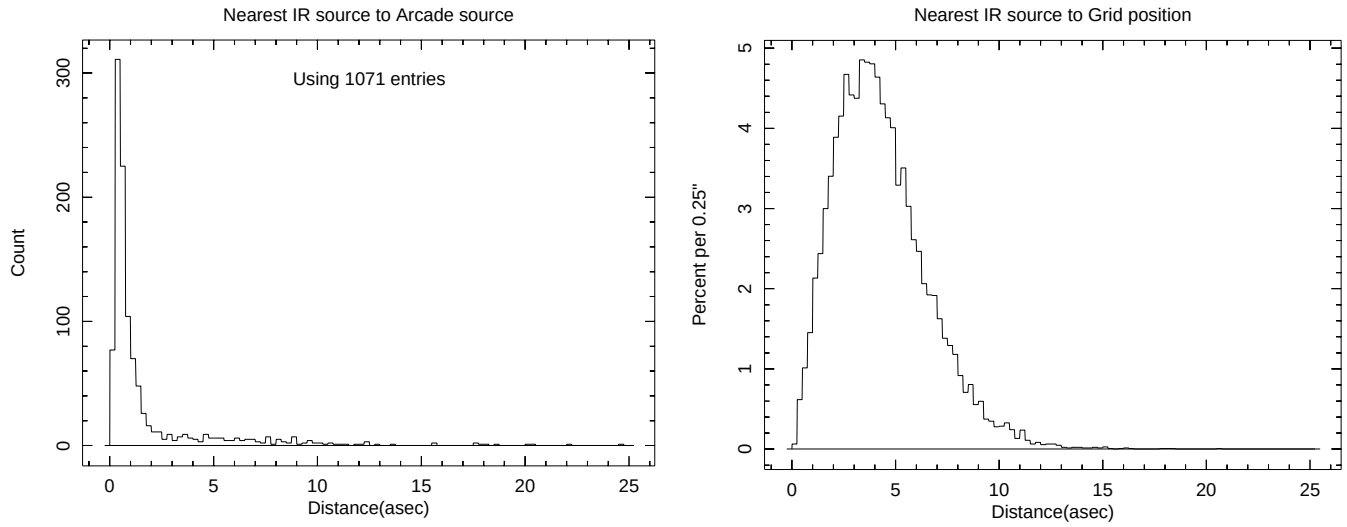

Fig. 3. Histogram of the distance to the nearest IR source for entries in a radio catalog selected by a target FDR of 5% (left) and to vertices of a 120×120 grid of positions separated by 10" (right).

TABLE I
RADIO SOURCES WITH IR MATCHES WITHIN 2"

Target FDR	SNR	match	no match	total	fract	FDR
5%	> 3	882	189	1124	0.176	7.7%
5%	> 4	732	115	888	0.136	0.5%
5%	> 5	546	69	645	0.112	0.2%
5%	> 7	363	45	431	0.110	$\equiv 0$

V. DISCUSSION

A technique of selecting a source catalog from an image using a estimate of the false detection rate based on the distribution of negative pixel values is described. An implementation in Obit is used on a deep VLA 2-4 GHz image. Available deep IR images and catalog were used to estimate the reality of entries in radio catalogs generated using a number of criteria. Table I includes a comparison of the IR detection rate and estimated false detection rate of catalog entries with various selection criteria including a pure 5% FDR criteria and a variety of additional SNR levels. The estimate of the achieved false detection rate for the target FDR 5% sample was 7.7%; however the less aggressive SNR> 4 sample had an estimated 0.5% achieved false detection rate. The technique seems to have produced nearly the desired results and allows the inclusion of fainter sources in the catalog produced. Use of the FDR target of 5% resulted in an increase of 74% in the number of entries in the catalog derived over the traditional SNR> 5 criteria. Of the additional sources, 86% are expected to be real.

ACKNOWLEDGMENT

I would like to thank Mark Lacy for access to the deep Spitzer images.

REFERENCES

[1] A. M. Hopkins, C. J. Miller, A. J. Connolly, C. Genovese, R. C. Nichol, and L. Wasserman, "A New Source Detection Algorithm Using the False-Discovery Rate," *Astron. J.*, vol. 123, pp. 1086–1094, Feb. 2002.
[2] D. A. Friedenberg and C. R. Genovese, "Straight to the Source: Detecting Aggregate Objects in Astronomical Images with Proper Error Control," *ArXiv e-prints*, Oct. 2009.
[3] W. D. Cotton and W. Peters, "False Detection Rate of Source Finding," *Obit Development Memo*, no. 25, 2011. [Online]. Available: ftp://ftp.cv.nrao.edu/NRAO-staff/bcotton/FDR.pdf
[4] J. C. Maduit and et al., *PASP*, vol. 124, p. 714, 2012.
[5] W. D. Cotton, "Obit: A Development Environment for Astronomical Algorithms," *PASP*, vol. 120, pp. 439–448, 2008.

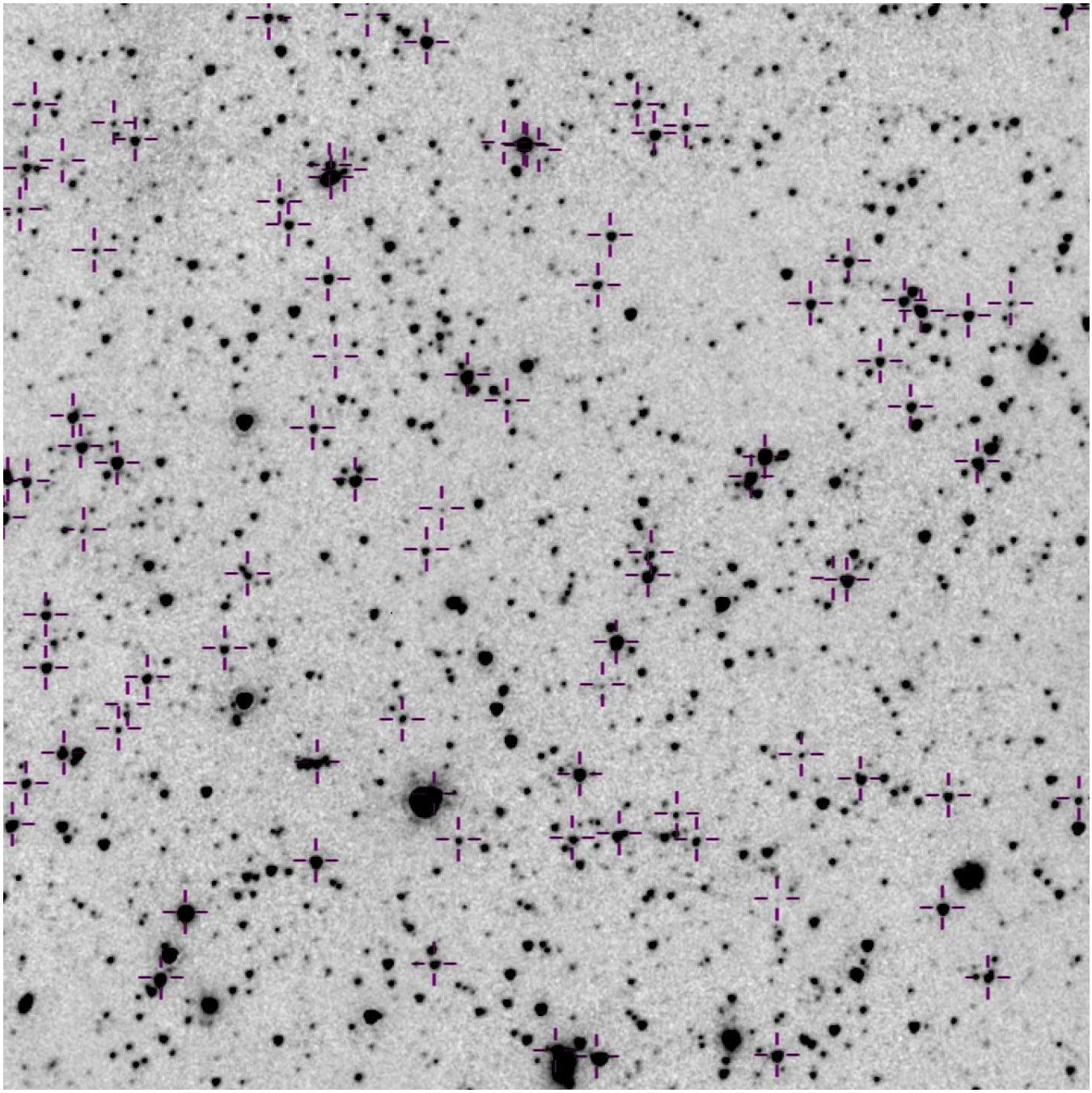

Fig. 4. Radio positions marked on negative gray scale of the 4.5 micron image.