Bull. Spec. Astrophys. Obs., 1996, 39, 155–160

On the properties of the extremal-median detector of faint radio sources of the indefinite shape

O.V. Verkhodanov, V.L. Gorokhov

Special Astrophysical Observatory of the Russian AS, Nizhnij Arkhyz 357147, Russia

Received December 4, 1995; accepted December 12, 1995.

Abstract. A faint signal detector constructed on the basis of the extremal and non–parametric statistics is described. The procedure of the detector efficiency estimation on the basis of false alarm and true detection probabilities is proposed. Results of the detector performance with the real RATAN–600 observational data are given. Comparison with other signal detectors is given.

Key words: statistics – data processing – signal detection

1. Introduction

The detection of extremely faint radio sources of indefinite shape is an important problem to be solved in radio astronomy (Davis, 1967, Parijskij, Korolkov, 1986, Mingaliev et al., 1991). Taking into consideration the fact that sources are faint and "sinking" in noises, their a priory shape is very difficult (or impossible) to describe with the beam pattern (or the point spread function). In these circumstances the employment of gaussian inscribing methods and correlation detection methods does not yield positive effect. And the data processing procedures lose their efficiency and even become inapplicable. The situation is complicated still more by a fact of the presence of the pulsed interferences and radio emission of the atmosphere.

The synchronous change in the stochastic properties in "columns" of the resolution intervals being examined is an important informative indication of a useful signal (Gorokhov, Verkhodanov, 1993, 1995). The sizes of these intervals are also the useful a priori information which is at user disposal.

The initial premises enumerated above allow formulation of a new statistical model of a radio astronomical signal.

2. Algorithm model

Let us have a series of synchronous scans $\{x_{ij}\}$. In the frames of this model these scans are looked through by a column of examined intervals $\{x_{i+t}\}_j$. Meanings of pixels are to be sporadic values with unknown distribution. The presence of a useful signal is interpreted as synchronous stochastic increasing (Leman, 1964, Gorokhov, Prokofiev, 1982) of pixels in the intervals in relation to pixels of the parts of the intervals where a useful signal is either lacking or extremely weak. The presence of interference is interpreted here as non–synchronous extreme values in the intervals. Part of interference effects can be removed here by the appropriate choice of sizes of the intervals being examined.

This model allows the modern achievements of non–parametric statistics to be used. In particular, we suggest makig use of extremal statisticses in developing the algorithm. These statisticses ensure the stability of the algorithm under the conditions when the type of pixels distributions is unknown and is changed. Extremal statisticses have an important property of immunity to the presence of correlation properties of interfering pixels. Pulsed interferences are removed with including the median statistics in the algorithm. To achieve invariance of thresholds, the relations of statistics are put in the procedure. Generalization of the model and algorithms to the two-dimensional case $\{x_{ij}\}_s$ (image reduction) is possible. Sampling of pixels, which do not contain a useful signal, using the extremal statistic (minimum of value), is an important constructive innovation. Intervals are supposed to be moved with a shift of interval value. This provides independence of samples and speeds up data processing. The effect of signal "splitting" is removed by the reverse motion with a shift of the initial interval by its halflength.

The modern achievements in mathematical statistics (asymptotical optimality of ranging procedures, methods of acceleration of machine tests) allow us to accompany the developed algorithms by quantitative characteristics of quality. They include the dependence of the probability of false alarms on the detection threshold, the dependence of the probability of true detection on the signal/noise ratio and on the probability of false alarms. These dependences allow an observer to estimate objectively results of detection of sources and to apply these estimates to forming the concept of survey completeness (Gorokhov, Verkhodanov, 1994).

Hereafter the detection algorithms realized in programs and results of simulated and natural tests of these algorithms are described in details.

3. Algorithm realization and its efficiency estimation

As as has been said above, authors propose extremal-median signal detector (EMSD) in a series of synchronous scans.

For this algorithm to operated several statisticses are calculated:

$$z_{ij} = \min_t\{\mathbf{D}_{i,j+t}\}, \qquad (1)$$

$$v_{ij} = \max_t\{\mathbf{D}_{i,j+t}\}, i = 1, m; j = 1, n; t = 1, r; \quad (2)$$

where $\mathbf{D}$ is the matrix of m vectors, m is the number of scans, n is the number of points in a scan, r is the number of points in a signal search interval. Using these two extremal statisticses the medians and the following statisticses are computed in each interval:

$$\theta_{v_j} = \operatorname*{med}_i v_{ij} - \operatorname*{med}_i z_{ij}, \qquad (3)$$

$$\theta_{z_j} = \operatorname*{med}_i z_{ij}, \qquad i = 1, m; j = 1, n; \qquad (4)$$

then their ratio is used as a test statistics:

$$q_j = \theta_{v_j}/\theta_{z_j}, \qquad j = 1, n \qquad (5)$$

In this case the hypothesis of object detection is adapted when the quantity q_j exceeds a certain quantity **c**. The threshold **c** can be either set by user or computed in the program.

It should be taken into the consideration in operations with the similar algorithms that all synchronous scans should have an equal zero level. Otherwise the low frequency background variations can produce an erroneous result, though the background fluctuations may appear to be interesting in the atmosphere influence estimation.

The algorithm efficiency was tested with model and also with real observational data.

The simulation was done using the RATAN–600 flexible astronomical data processing system FADPS (Verkhodanov et al., 1993). The recordings are simulated by two types of five scans each: pure noises (with non–correlated gaussian noise) and noise plus signal (a source with intensity gaussian distribution of a certain amplitude and halfwidth is put in each interval).

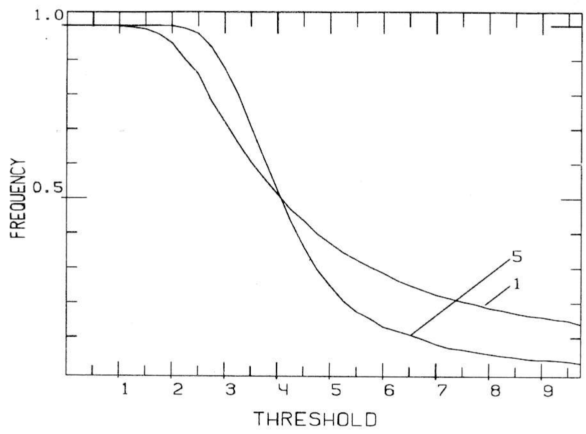

Figure 1: *False alarm probability vs detection threshold for 1 and 5 scans in EMSD operation.*

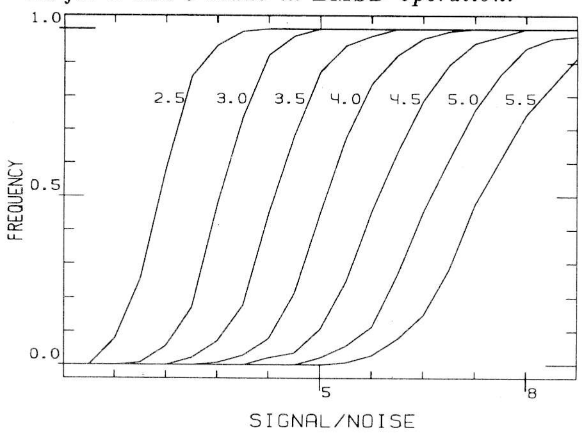

Figure 2: *The true detection probability as a function of the signal/noise ratio for different detection thresholds in EMSD operation.*

The detector was passed over the pure noise recordings to compute the probability of false alarm, and the number of events was counted in depending on detection threshold.

The results of computation of the false alarm probability as dependent on the detection threshold for one and five synchronous scans are shown in Fig. 1.

To compute the true detection probability depending on the signal/noise ratio, the authors simulated recordings with different signal/noise ratio in five scans for each ratio. The detector with the a specified set detection threshold was passed by these recordings. The family of the curves, computed for five scans and different detection thresholds, is shown in Fig. 2.

The optimum choice of the threshold consists in selection a value, which, on one hand, yields a low level of false alarms, and on the other hand, yields a sufficiently large probability of true detection for a

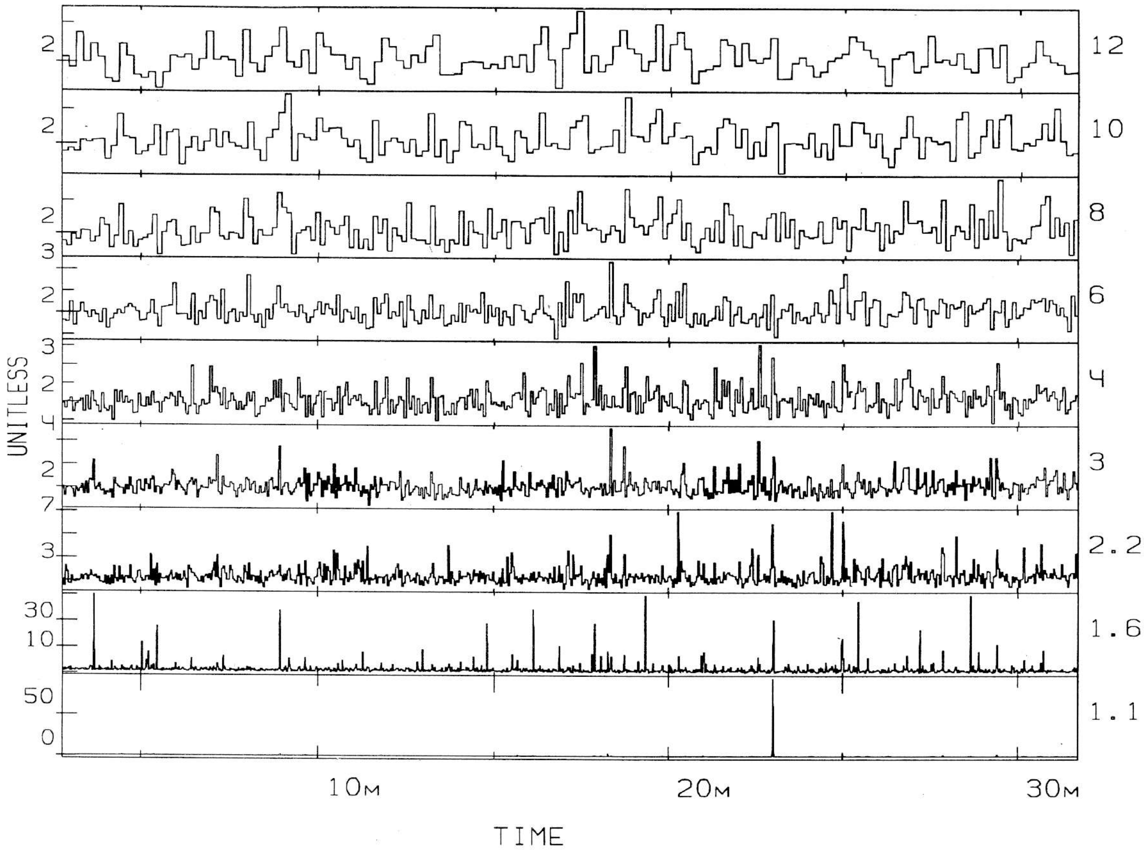

Figure 3: *Results of EMSD operation dependening on the interval size for real data recordings. The interval sizes are shown at the right side. A size of beam pattern is 2.2 sec. The ratio (5) is plotted as ordinate. See also Fig. 5.*

small signal/noise ratio. As shown in Fig. 1 and 2, this value is approximately equal to 3 for extremal median signal detection.

It should be noted that this signal detector can operate for single recordings also. But the main gain still consists in the use of synchronism of several recordings.

One can fast assess quantatively parameters of the found source after the signal detection in any interval (when the statistics q_j exceeds the input threshold c using the following formulas (Fomalont, 1989):

$$I = \sum_t I_t$$

$$X = \frac{1}{I} \sum_t X_t I_t \quad (6)$$

$$B = \sqrt{\frac{1}{I}(X_t - X)^2 I_t}, \qquad t = 1, r;$$

where I_t is the intensity value in the t-th point of the interval, X_t is the position of this point in the recording, I is the integral intensity of the possible object, B is the size of the object, r is the size of an interval in pixels.

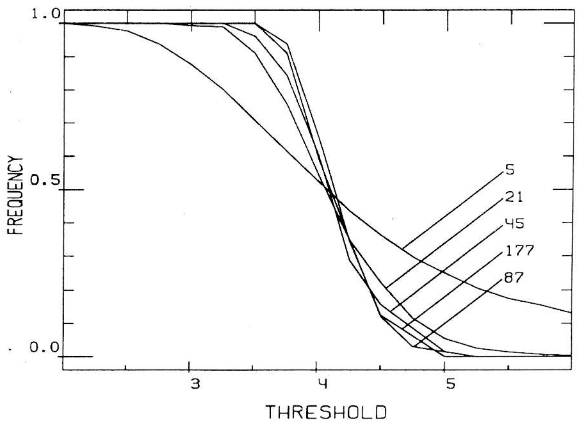

Figure 4: *Increase of the probability of false alarm after decrease of the interval size in the detection thresholds of 2.5 – 5 for EMSD. Numbers show the interval sizes for the pointed curves.*

Parameters of the object can also be estimated with the Gauss-analysis procedure (Ivanov, 1979). The time of this procedure execition is sharply decreased with decreasing interval. The new algorithm

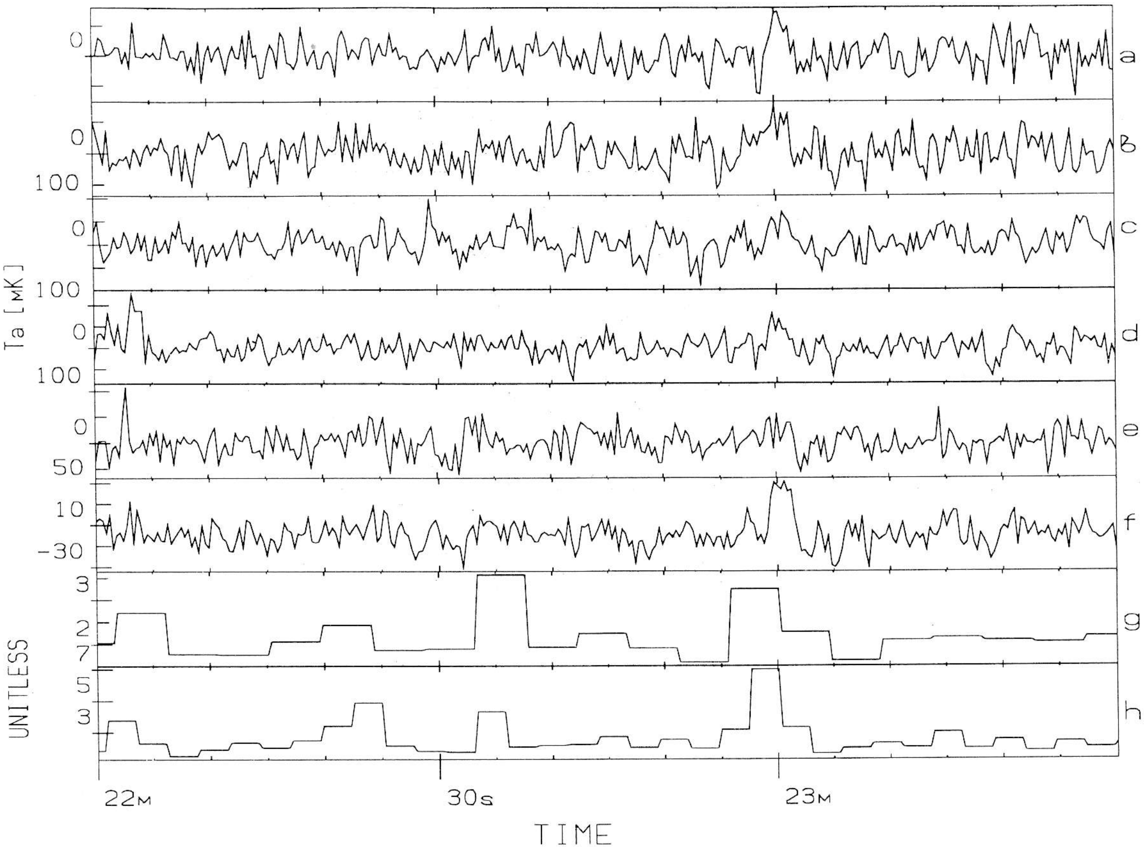

Figure 5: *Results of EMSD operation for the different input intervals: g – interval of two beam patterns, h – one beam pattern, for real data of five synchronous scans (a–e). To compare results the scan with robust average of these data is shown (f). The antenna tempretature is for the first six scans, and ratio (5) is for the two last scans plotted along the Y–axis.*

can help to do it.

As has been said, an essential thing for statistical detectors is the choice of a signal searching interval. This interval is defined, first of all, by the size of a signal to be found. Thus knowledge of the additional a priory information about apparatus function may play a great role.

When one operates with an interval less than the size of the possible desired object, the statistics q_j may sharply increase at the moment of calculations at the object boundary (see Fig. 3).

However, the probability of false alarm grows also for due to decreasing amount of the statistical data (see Fig. 4).

In this case it is impossible to estimate the characteristics of the object. The best interval from the authors' view point is an interval with the size 1.2–1.5 that of the object. Results of EMSD operation on real RATAN-600 observational recordings are shown in Fig. 5.

4. Comparison with other detectors

It is interesting to compare qualitatively the EMSD with detectors of other types, e.g. integral or correlational ones.

The integral detector can be constructed changed by substituting the sum of pixels in an interval fro the right side of formula (3):

$$\theta_{v_j} = \sum_i \sum_t \mathbf{D}_{i,j+t}, \qquad (7)$$

i.e. we have the boxcar average. Formula (4) for the statistics θ_{z_j} can be left unchanged.

For the correlational detector we have

$$\mathbf{C}_{ij} = \frac{1}{A} \sum_s \mathbf{D}_{i,j-s} A_s \qquad (8)$$

Here A is the integral by the beam pattern (apparatus function) A_s, v_{ij} and θ_{v_j} are defined by the following formulas:

$$v_{ij} = \max_t \{\mathbf{C}_{i,j+t}\}, i = 1, m; j = 1, n; t = 1, r; \qquad (9)$$

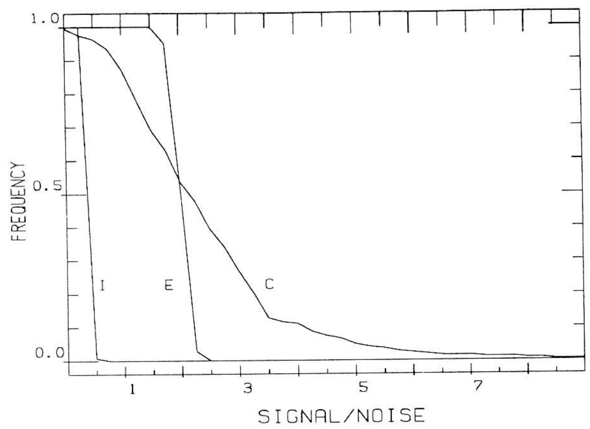

Figure 6: *False alarm probabilities vs the detection threshold for extremal-median (E), integral (I) and correlational (C) detectors.*

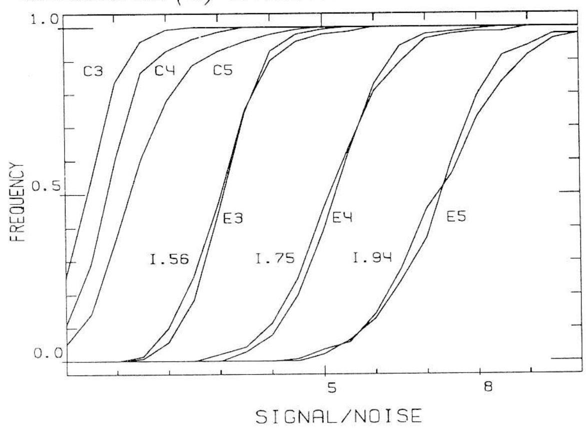

Figure 7: *Relationships between the true detection probabilities and the signal/noise ratios for the different detectors and detection thresholds. Symbols E, C and I show detector types (correspondingly EMSD, correlational and integral), numbers at the right of symbols show the detection thresholds.*

$$\theta_{v_j} = \operatorname*{med}_i v_{ij} \qquad \text{or}$$

$$\theta_{v_j} = \frac{1}{m} \sum_i v_{ij}, i = 1, m; j = 1, n. \qquad (10)$$

One can apply formula (4) to the statistics θ_{z_j} for a correlational filter or simply calculate dispersion in the current interval. False alarm probabilities versus the detection threshold for extremal-median, integral and correlational detectors are shown in Fig. 6.

The relationships between the true detection and the signal/noise ratio for different detection thresholds are shown in Fig. 7 for all the three filters.

As it is shown in Fig. 7 we have a very low probability of the false alarm even for threshold 1 in the case of an integral filter. And it is necessary to set a threshold practically an order less than for the extremal median filter for obtaining sufficient probability for the object detection. The probability of signal "smearing" is rather high for the integral filter when positive peak is added to the negative noise values and can practically be completely neutralized.

For the correlational signal the probability of true detection is very high for small signal/noise ratios. But the probability of false alarm is also very large for the sufficiently large detection thresholds.

5. Other detectors built on the extremal statistics

The procedures described above allow us to build and study other signal detectors operating in a series of synchronous scans. For example, one can construct a detector responding to signal changing (variability) in the series of synchronous scans. The statistics z_{ij} and v_{ij} are defined by formulas (1) and (2). The statistics θ_{v_j} is defined by the following formula:

$$\theta_{v_j} = \max_i v_{ij} - \min_i v_{ij}, \qquad (11)$$

The statistics θ_{z_j} can be defined by formula (4).

Extremal statistics can be also applied to detection of "spotness" or image segmentation (Gorokhov, Bursian, 1990). Here the statistics θ_{v_j} is defined by formula (11), and the statistics θ_{z_j} is computed by the following formula:

$$\theta_{z_j} = \max_i z_{ij} - \min_i z_{ij}. \qquad (12)$$

References

Davis, M.M.: 1967, Bull. Astr. Inst. Netherlands, **19**, 201.

Formalont, E.B.: 1989, in: Perly, R.A, Schwab, F.R., Bridle, A.H. (eds.), A Collection of Lectures from the Third NRAO Synthesis Imaging Summer School, Astronomical Society of Pacific, 213.

Gorokhov, V.L., Prokofiev, V.N.: 1982, Avtometria, Leningrad, **6**, 36.

Gorokhov, V.L., Bursian, A.A.: 1990, Izvestiya of Leningrad Electro-Technical Institute, **426**, 36.

Gorokhov, V.L., Verkhodanov. O.V.: 1993, Prepr. SAO # 89SPb. St.Petersburg branch.

Gorokhov, V.L., Verkhodanov, O.V.: 1994, Pis'ma Astron. Zh., **20**, 776.

Gorokhov, V.L, Verkhodanov, O.V.: 1995, in: Reports thesis of the 26th radio astronomical Conference. St.Petersburg, Institute of Applied Astronomy of RAS. 51.

Ivanov, L.N.: 1979, Astrofiz. Issled. (Izv. SAO), **11**, 213.

Leman, E.: 1964. Proverka statisticheskih gipotez (The Check of Statistical Hypotheses). Moscow, Nauka.

Mingaliev, M.G., Verkhodanov, O.V., Khabrakhmanov. A.R., Golnev, V.Ya., Temirova, A.V.: 1991, Soobshch. Spets. Astrofiz. Obs., **68**, 47.

Parijskij, Yu. N., Korolkov, D.V.: 1986, Experiment Cold: the first deep survey with the RATAN–600 radio telescope, in: Syunyaev, R.A. (ed.), Sov.Sci.Rev.E. Astrophys. Space Phys., Harwood Academic Publishers GmbH, Switzeland, 5, 39.

Verkhodanov, O.V., Erukhimov, B.L., Monosov, M.L., Chernenkov, V.N., Shergin, V.S.: 1993, Bull. Spec. Astrophys. Obs., **36**, 132.