Available online at www.sciencedirect.com
SCIENCE DIRECT
PERGAMON
Acta Astronautica 58 (2006) 222–229
ACTA ASTRONAUTICA
www.elsevier.com/locate/actaastro

SETI-Italia 2003 status report and first results of a KL transform algorithm for ETI signal detection

S. Montebugnoli[a,*,1], C. Bortolotti[a], A. Cattani[a], A. Maccaferri[a], A. Orlati[a], M. Poloni[a], S. Poppi[a], J. Monari[a], M. Roma[a], P.P. Pari[a], M. Teodorani[a], S. Righini[b], C. Maccone[c], D. Caliendo[d], C.B. Cosmovici[e], N. D'amico[f]

[a] *Institute of Radioastronomy, INAF, Via Gobetti 101, 401129 Bologna*
[b] *Osservatorio Astronomico di Torino, Via Osservatorio 20, Pino Torinese*
[c] *International Academy of Astronautics, Via Martorelli 43, 10155 Torino (Italy)*
[d] *Under graduating Student at the Institute of Radioastronomy*
[e] *Istituto di Fisica dello Spazio Interplanetario CNR*
[f] *Osservatorio Astronomico di Cagliari Loc. Poggio dei Pini, 09012 Capoterra (CA)*

Received 21 June 2005; received in revised form 31 August 2005; accepted 13 September 2005
Available online 28 November 2005

Abstract

SETI-Italia is the Italian national SETI Program. It is run as a part of the Italian Bioastronomy Project, by IRA (Istituto di Radioastronomia) a subsidiary of the Istituto Nazionale di Astrofisica. At the moment, Italy is the only European country conducting a continuous SETI "listening" program, while occasional SETI searches may be conducted in France and the Ukraine. Outside Europe, continuous SETI programs are conducted in the USA, Australia and Argentina. The SETI-Italia activities started early in 1998 with a Serendip IV system connected in piggy-back mode to the Medicina VLBI 32-m antenna dish. In the years 2002/2003, considerable efforts were devoted to set up a fast computational system capable of computing the KLT (Karhunen–Loève transform). This is a virtually new mathematical procedure in SETI, and is much more general than the FFT. It can search for the presence of signals both narrow-band and wide-band embedded in noise both coloured and white. The mathematical features the KLT have already been described in this Conference by Claudio Maccone [Innovative SETI by the KLT, Pešek Lecture 2003, paper # IAA.9.1.01 presented in Bremen, in this 2003 International Astronautical Congress [1]; Telecommunications, KLT and Relativity, vol. 1, IPI Press, Colorado Springs, CO, 1994, ISBN #1-880930-04-8 [2].]. Our software implementation of the KLT follows these lines and also is indebted to some basic ideas put forward by Robert S. Dixon [On the detection of unknown signals, in: G. Seth Shostak (Ed.), Third Decennial US–USSR Conference on SETI, Astronomical Society of the Pacific Conference Series, vol. 47, 1993, pp. 128–140] of the Ohio State Radio Observatory as early as 1993.
© 2005 Elsevier Ltd. All rights reserved.

1. SETI-Italia

SETI (an acronym for Search for ExtraTerrestrial Intelligence) is a research program currently pursued in several technologically advanced countries (the USA, Italy, Argentina and Australia, and, occasionally, France and the Ukraine) (see footnote 1). SETI is conducted by virtue of the most modern and powerful

* Corresponding author.
E-mail addresses: Stelio@ira.cnr.it (S. Montebugnoli), dimartino@to.astro.it (S. Righini), clmaccon@libero.it (C. Maccone).

[1] IAC-03-IAA.9.1.02 paper presented at the 54th IAF Congress in Bremen (2003) by S. Montebugnoli.

0094-5765/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.actaastro.2005.09.003

Fig. 1. The Medicina VLBI dish and, in the forefront, part of the Northern Cross Array (North/South arm).

radiotelescopes that search the sky for radio signals either monochromatic (i.e. "intentional") or modulated (i.e. "possibly unintentional" = "leakage") arriving from extraterrestrial technologically advanced civilisations. Because of the rotation of both planets (the Earth and the Exoplanet), every such radio signal is expected to be frequency shifted (Doppler effect).

By using the Medicina VLBI 32 m dish, we started in 1998 a systematic search conducted in parallel to the ongoing standard activities ("piggy-back" mode). We used a high-resolution spectrum analyser named SERENDIP IV (see Ref. [3]) in a 4 million channel configuration.

Thanks to substantial support by the SETI Institute, in 1999 the system was upgraded to 24 million channels at 15 MHz input bandwidth. Up to now, observations have been carried out in the radio astronomical bands included in the 1.4–23.5 GHz range. This allows the programme to exploit the 100% of the antenna allocation time at extremely low cost, while taking the RFI (= Radio Frequency Interference) scenario continuously under control in parallel with the Sentinel 2 RFI monitoring control system 3.

2. The instruments

The Istituto di Radioastronomia (IRA) of the Istituto Nazionale di Astrofisica (INAF) (formerly Italian National Research Council = CNR) runs 3 radiotelescopes:

1. The Medicina (near Bologna) 32 m VLBI dish shown in Fig. 1. Operating frequency: 1.4–23.5 GHz.
2. The Noto (near Siracusa) 32 m VLBI dish, shown in Fig. 2, equipped with active mirror. Operating frequency: 1.4–43 GHz.

Fig. 2. The Noto 32 m VLBI dish, equipped with an active mirror ($F_{max} > 33$ GHz).

Fig. 3. The large Northern Cross Array, 560 m (East–West arm) ×640 m (North–South arm).

3. The 30, 000 m^2 collecting area Northern Cross Array, a large cylindrical reflector T-shaped array. Operating frequency: 408 ± 2.5 MHz (Fig. 3).

The Northern Cross Array (5,632 dipoles) is one of the largest transit telescopes in the north hemisphere. Its collecting area is equivalent to that of 3 football

stadiums) and, for this reason, it is one of the most suitable instruments to be transformed in a very large square kilometre array (SKA) test bed. Finally, a 64 m class parabolic antenna is currently under construction at San Basilio, close to Cagliari (Sardinia island).

3. Activities

The SETI-Italia observations are currently ongoing by virtue of a 24 million channel Serendip IV high-resolution spectrometer, connected in piggyback mode to the 32 m VLBI dish located at the Medicina station. A post-processing algorithm for pattern recognition of the Doppler-shifted signal, based on the Hough transform (= Radon Transform for a straight line), has been successfully tested. Operating with this very efficient transform, many monochromatic candidates (Dopplered signals) have been detected in the 1.4/1.6 GHz band, but none of them have been further confirmed with follow up observations. Principal bands observed during the normal telescope activities have been,

- 1.4–1.6 GHz,
- 4.8–6.5 GHz,
- 8.5 GHz,
- 22–23.5 GHz,

within the European VLBI observation run and single dish antenna programs (Blazar, Geodynamic and Spectroscopy).

4. The KLT

Until recently, the basic SETI paradigm was

(1) Suppose ET is sending a radio signal out into space saying something like "Hi, here we are!"
(2) Then, this signal needs to be quite simple in order to be detected and distinguished from the natural wide-band radio signals coming from space. In other words, the ETI signal must be a CW (continuous wave) signal, that is a radio carrier.
(3) The most suitable mathematical detection tool to detect such a simple, sinusoidal radio carrier is the high-resolution spectrum analysis based on the fast Fourier transform (FFT). In fact, this transform assumes the signal to be periodic with a period equal to the acquisition time (which is not true in general) and tries to reconstruct signals of any kind by virtue of sines and cosines only (orthonormal base functions). It turns then out that, while the FFT works very properly with monochromatic signals, it starts losing its efficiency in case of wide-band modulated signals.
(4) A more general transform, able to "understand" whether or not a radio band contains any kind of modulated signals (including non-periodic ones), is the Karhunen–Loève transform (KLT), 1946. This transform appears to be a very promising ETI signal detection tool because it extracts the transformation base functions from the signal itself. If an extraterrestrial civilisation exists and transmits radio signals (intentionally or unintentionally), we absolutely do not know which kind of modulated or not modulated signals are radiated into space. As already mentioned, we presume that they are sending a CW signal, but this is only an assumption. For this reason we need a very efficient transform able to recognise any kind of complex modulated radio signals embedded in the noise. Such a transform is the KLT. Unfortunately it requires a very high computational load from the processing engine. In practice it involves the calculation of the eigenvectors and eigenvalues of very large autocorrelation matrices. The KLT rests on the fact that, given any observation $X(t)$, it is possible to separate the statistical component Z_n from the temporal component as follows

$$X(t) = \sum_{n=1}^{\infty} Z_n \phi_n(t),$$

where

(1) Z_n is the random variables not changing in time (and similar to the coefficients of the sines and cosines in any Fourier series), and
(2) $\phi_n(t)$ the base functions, i.e. a set of orthonormal basis functions that turn out to be the eigenfunctions (or eigenvectors in the Hilbert space of the square-integrable functions) of the autocorrelation of $X(t)$.

The KLT represents an optimal transform, from the mean square point of view, with some peculiar characteristics

$\langle Z_n \rangle = 0$, mean value of the coefficients Z_n,
$\langle Z_n Z_m \rangle = 0$, statistical independence from the autocorrelation point of view,
$\sigma^2_{Z_n} = \lambda_n$, Z_n coefficients variance,
$\int_0^T \phi_m(t)\phi_n(t)\,\mathrm{d}t = \delta_{mn} = \begin{cases} 0 & \text{for } m \neq n, \\ 1 & \text{for } m = n, \end{cases}$
(orthonormality of the base functions).

The algorithm for the KLT is, in practice, the solution of the eigenvectors/eigenvalues computation problem. The eigenvalue of a matrix A (i.e. the autocorrelation matrix) is a scalar to which corresponds an eigenvector different from zero, leading to the expression

$$Ax = \lambda x.$$

For a matrix A of order N, we have N eigenvectors and N eigenvalues. Due to its extremely high computation complexity and to the fact that it cannot be distributed (parallelized), at present this transform cannot be computed in real time. On the other hand it is a detection tool and then the real time is not really required. We only need a sufficient number of "checks" per second (sampling in time) to know whether or not an ET signal is present in the received band. For these reasons in the beginning our efforts have been aimed to get the maximum speed but not the real time. In this way the following assumptions have been made:

- After the acquisition of N points, it is sufficient to compute only the first M (dominant) eigenvalues because the $N - M$ eigenvectors, related to the noise, are not important.
- Since the algorithm for the eigenvalues computation is iterative, accepting a certain approximation in their determination makes such a task much faster. On the other hand, we only need to know a good estimation that allows us to sort these dominant eigenvalues in a decreasing order.

5. The algorithm

With the above mentioned assumptions, the IRA eigenvalues algorithm for the estimation of the eigenvalues of the autocorrelation matrix, is based on the following steps (Fig. 4):

N points data acquisition: N points are acquired via a 12 bit A/D converter and sent to the computation block via a FPDP wideband bus.

N points autocorrelation vector computation: The N components data vector are used to compute the $2(N-1)$ components autocorrelation vector (symmetric in $N-1$)

Autocorrelation matrix computation: The $(N \times N)$ autocorrelation matrix (here $N = 5$)

$$A_{\text{matrix}} = \begin{bmatrix} a & b & c & d & e \\ b & a & b & c & d \\ c & b & a & b & c \\ d & c & b & a & b \\ e & d & c & b & a \end{bmatrix}$$

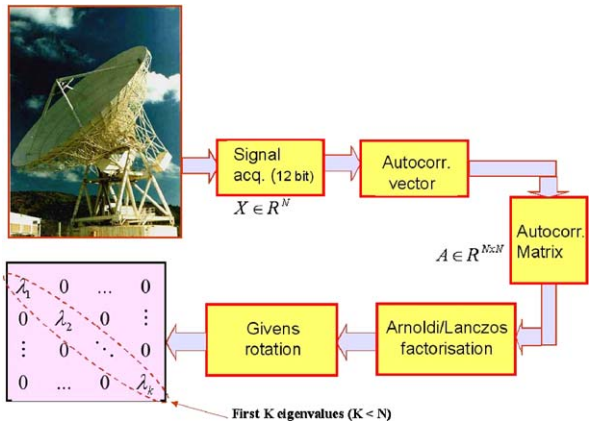

Fig. 4. Block diagram of the KLT algorithm developed by IRA (Istituto di Radio Astronomia) for SETI-Italia.

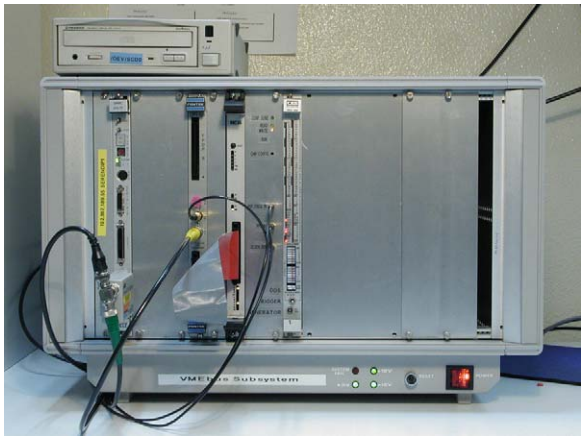

Fig. 5. The ALTIVEC^{TM} 7400 vector multinode system.

is computed by starting from the autocorrelator vector CORR vector $= [e\ d\ c\ b\ a\ b\ c\ d\ e]$.

This is a Toeplitz matrix that contains the energy, given by the autocorrelation estimator, on the main diagonal.

Arnoldi/Lanczos factorisation: The $(N \times N)$ dimension of the A_{matrix} are drastically reduced using the Arnoldi/Lanczos factorisation. This produces a K order $(K \ll N)$ three-diagonals square matrix (L matrix), whose eigenvalues computation turns out to be much easier and faster.

Givens Rotation (Eigenvalues approximation): The Givens rotation algorithm is used to extract the eigenvalues of the L_{matrix}. The output of this block gives a diagonal K order matrix containing (ranked by a decreasing order) a good approximation of the A_{matrix} dominant eigenvalues, i.e. the first K eigenvalues.

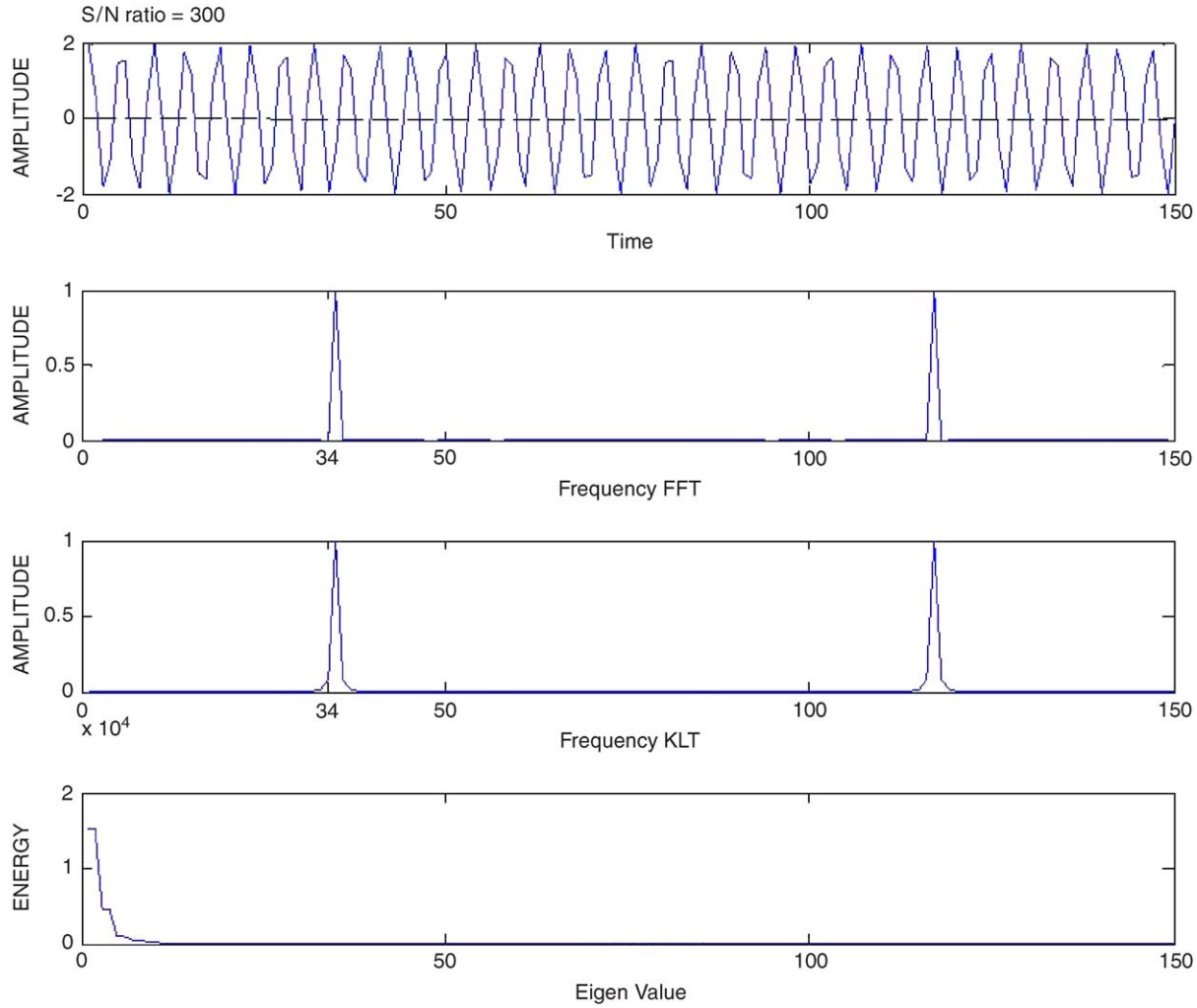

Fig. 6. FFT and KLT of a signal with $S/N = 300$.

A 12 bit, 40 MS/s Pentec® A/D converter and a Mercury® MCJ6-PPC7400 double CPUs boards, operating in a VME environment (Fig. 5), are in charge to run all the above mentioned steps of the algorithms. Such a system is modular, meaning that up to several hundreds of CPUs can be parallelised to increase the computational speed and/or bandwidth.

The performances of three different algorithms for the computation of the ten dominant eigenvalues (5 MHz of input bandwidth and $N = 1024$ points) have been tested on the Mercury® ALTIVEC board, yielding the following results:

Jacobi algorithm	4158 ms,
Lanczos alg. (standard)	67.5 ms,
IRA procedure	6 ms.

6. Signal analysis

A 34 Hz signal with different levels of injected noise has been used in a simulation to demonstrate the behaviour of the FFT and KLT. In Fig. 6, a $S/N = 300$ signal is represented. Here, and in the following Figs. 7 and 8, the first graph shows the function in the time domain, the second its FFT, the third is the FFT of the dominant eigenvector. The last graph is the plot of the eigenvalues ranked by decreasing amplitude (eigen spectrum).

With such a S/N, no difference between the FFT and KLT are noted (second and third graph). Introducing a certain amount of noise so as to obtain a $S/N = 1.0124$, the different behaviour of both the FFT and KLT becomes evident, as reported in Fig. 7. The FFT (second graph in the same figure) is not able anymore

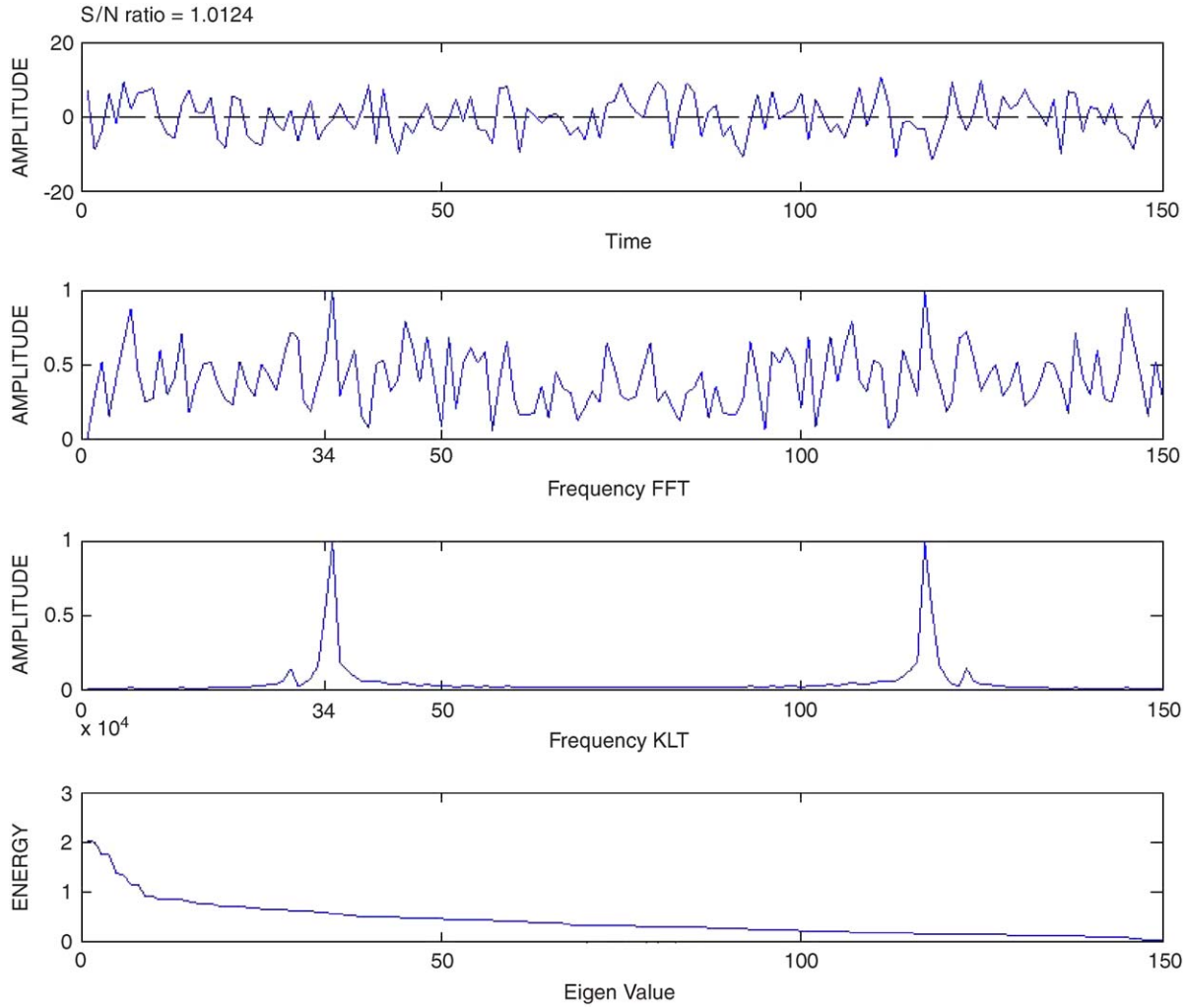

Fig. 7. FFT and KLT of a signal with $S/N = 1.0124$.

to detect the line while the KLT (third graph) is clearly able to detect it. Increasing again the injected amount of noise up to $S/N = 1.0073$, only the KLT is able to detect the signal as visible in Fig. 8 (third graph).

The basic work of Robert S. Dixon, Ref. [4], was focused on the use of the dominant eigenvalues as a detection key. This can work properly when the S/N ratio is high as in the case of Fig. 6. Here the major part of the energy (fourth graph of the same figure) is concentrated in the dominant eigenvectors and then it is easy to detect the presence of a signal by the analysis of the eigen spectrum. This detection method has problem in the detection of noisy signal as in reported in Fig. 8.

Here the energy is completely smeared on all the eigenvalues. It could then be difficult to detect a signal this way.

7. Our approach

Based on the above simulations, we implemented a first analysis method that seems to give good results. Of course this is a very preliminary approach and definitely needs more investigations.

Our method is based on the accumulation of the average of the dominant eigenvectors (axis) FFTs containing the best projection of the acquired cluster of data. There are four main steps:

- Computation of the dominant eigenvectors (say the first p eigenvector $\rightarrow$ axes).
- Computation of the FFTs of the p eigenvectors.
- Channel by channel average of the p spectra, in order to obtain a single spectrum P.

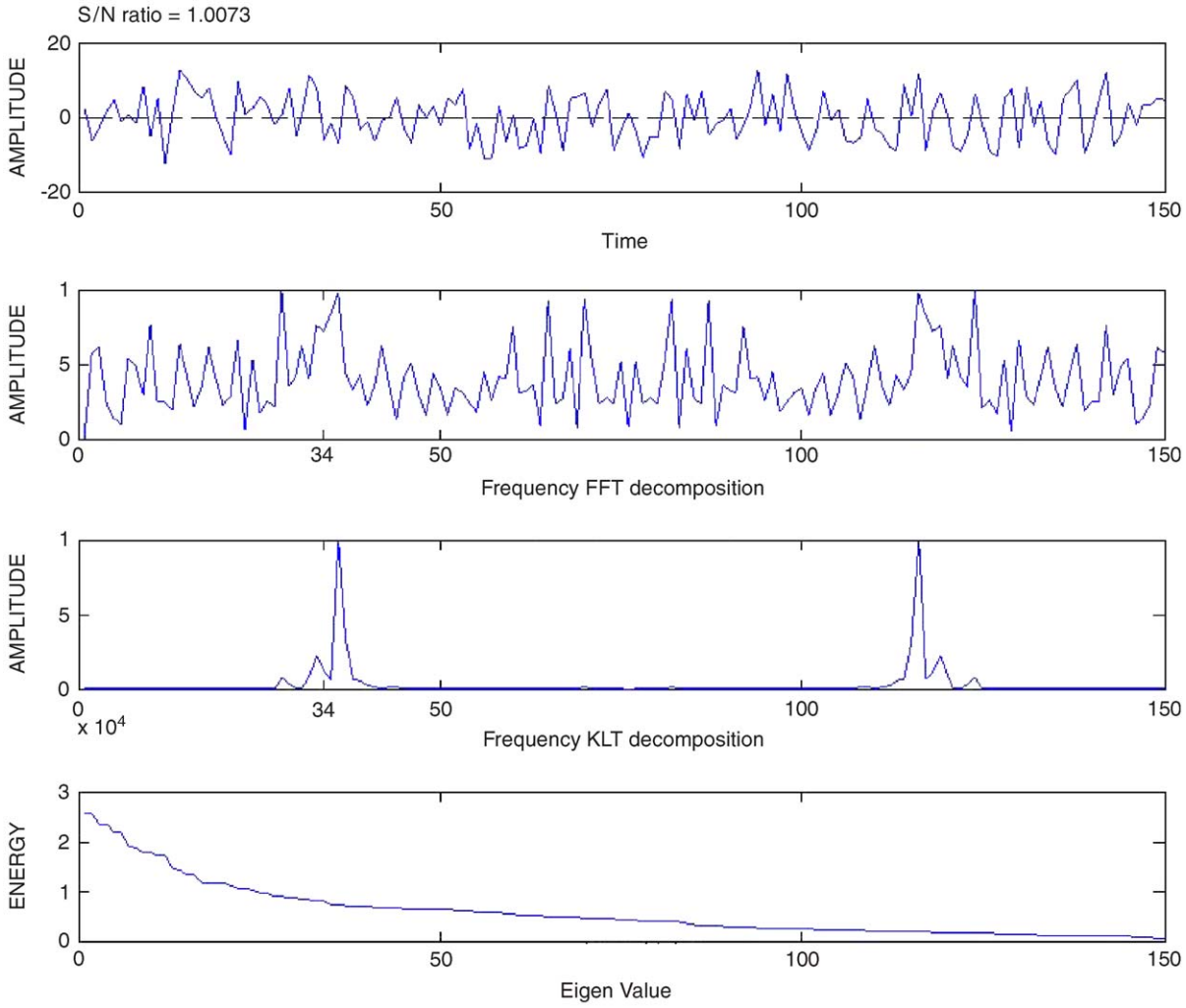

Fig. 8. FFT and KLT of a signal with $S/N = 1.0073$.

- Average of the necessary number of single spectra P to obtain the final spectrum with the desired rms noise.

To check the procedure, the system has been connected to the VLBI antenna and an OH maser (6.035 GHz) has been observed. To subtract the filter shape and Radio Frequency Interference (RFIs) if present, some observations have been performed on source and some other off source. A qualitative result. is reported in Fig. 9.

The total average time (exposure) of the W3 OH Maser is here equivalent to a few tens of seconds in the On source position.

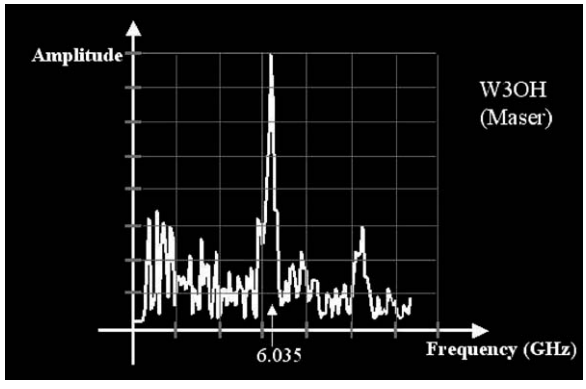

Fig. 9. Spectrum of the W_3 OH Maser.

8. Conclusion

Until the summer 2003, no evidence of ETI signals was obtained from SETI observations at the Medicina radiotelescopes. But the Serendip IV system is continuously operating for both SETI observations and RFI monitoring.

Preliminary tests were performed on the signal extraction from the noise by use of a KLT based algorithm. The approach we present in this paper is much faster than the other ones because we compute only the dominant eigenvectors/eigenvalues and, in addition, we accept to obtain only a good estimation of them, rather than computing their precise values. In our approach an average of the dominant eigenvectors FFTs, ready to be further averaged per each cycle, is considered. This is necessary to get a final spectrum with the requested rms noise. The sum of the averaged FFTs of the best data projection of the components on the dominant axis seems to give much more detection possibilities than using only the eigenvalues alone as a detection key, as thought up to now.

This method is a starting point for further investigations, and developments on this very powerful transform for signal detection in the SETI program.

References

[1] C. Maccone. Innovative SETI by the KLT, Pešek Lecture 2003, paper # IAA.9.1.01 presented in Bremen, in this 2003 International Astronautical Congress.

[2] C. Maccone. Telecommunications, KLT and Relativity, vol. 1, IPI Press, Colorado Springs, CO, 1994, ISBN #1-880930-04-8.

[3] D. Werthimer, et al., The Serendip IV Arecibo Sky Survey—a new Era in Bioastronomy, Astronomical Society of the Pacific Conference Series, vol. 213, 2000, pp. 479–483.

[4] R.S. Dixon, et al., On the detection of unknown signals, in: G. Seth Shostak (Ed.), Third Decennial US–USSR Conference on SETI, Astronomical Society of the Pacific Conference Series, vol. 47, 1993, pp. 128–140.