IAC-03-IAA.9.P.02

Generalized Hough Transform: a Useful Algorithm for Signal Path Detection

Author: J. Monari
E-Mail: jmonari@ira.bo.cnr.it
INAF- Istituto di Radioastronomia, PO Box 14 - Via Fiorentina, 40060 Villafontana,-Bo-
Phone: ++39 051 6965846,
Fax: ++39 051 6965810

Co-author: S. Montebugnoli
E-Mail: jmonari@ira.bo.cnr.it
INAF- Istituto di Radioastronomia, PO Box 14 - Via Fiorentina, 40060 Villafontana,-Bo-
Phone: ++39 051 6965827,
Fax: ++39 051 6965810

A. Orlati[1], M. Ferri[2], G. Leone[2].

1-INAF- Istituto di Radioastronomia, PO Box 14 - Via Fiorentina, 40060 Villafontana,-Bo-
2-Università di Bologna, ARCES "E. De Castro" and Mathematics Dept.

Abstract
How is it possible to recognize ETI signals coming from outer planets?
This is one of the several questions that SETI researchers must answer to.

During the first months of 1998, the Italian SETI program started in Medicina with the installation of the Serendip IV 24Million Channel digital spectrometer. This system acquires every day a huge quantity of data, which must be processed off line, in order to detect the possible ETI signals stored inside them.
This process has to be as automatic as possible, and it must filter any interference, detecting particular ETI paths and storing the results in a database where the operator will be able to check and verify them. In our group, several programs are devoted to this topic, and they are collectively called SALVE 2 (Software Aimed at Off Line Verification Eti ver 2.0).

The problem of automatic elaboration and signal recognition is very important now, and it will be felt even more in the future when a new instrument as the SV is ready. The heart of the whole process is the algorithm for ETI signal recognition. In fact to do that, longer before decoding possible modulated signals, it is necessary to isolate them from thousands of human interferences and natural emissions originated on our planet or elsewhere in the universe. It is generally thought that a possible ETI emission could be a monochromatic signal, dopplered by the relative motion of the planet with respect to the Earth.

Here is presented a natural evolution of a previous work, which was based on a simple Hough transform and was limited only to the detection of short linear tracks in the Join Time Frequency Matrix (JTFM) stored by SIV. The new generalized Hough algorithm allows to detect the sinusoidal tracks by the transformation of the JTF bidimensional Cartesian space (x,y), in the Generalized Hough quadridimensional space, where the main vectors are the sine parameters Amplitude, Frequency, Phase and Offset.

At the end of the paper some results, obtained with the computation of real and simulated JTFM, are shown.

§1 Introduction

New technologies in development aim to the realization of more sensitive and performing instruments. Designs of the newest future radiotelescope (SKA, Allen telescope) and back ends (Serendip V, MEDALT1) will allow us to deepen our radio-signals analysis, increasing the probability to detect a signal emitted by Extra Terrestrial Intelligence (ETI).
This will mean a huge amount of data to compute. The processing will be possible only in automatic mode, exploiting the computing power of the future computers and a proper algorithm describing "the type of signal we think it will be possible to receive".
Here is described an interesting algorithm used for SALVE 2 [2] data post processing, a collection of procedures written from the experience managed by our Institute after the installation of the multi channel spectrum analyzer Serendip IV (SIV [3]).
This paper is organized as follows: §2 describes the data set structure of JTFM matrix, §3 presents SALVE 2 programs, §4 From Radon to generalized Hough Transform, §5 presents the results and discusses future directions

§2 JTFM Matrix

It is very important to know the data set structure of the Join Time Frequency Matrix (JTFM) stored by SIV, to understand how it is possible to recognize an ETI signal.
We could consider the stored data as a matrix where the columns are the frequencies while the rows correspond to time. Each row represents the bandwidth of SIV (15MHz for the Italian instrument).
Each channel is "ON" only when a signal hits a particular user-defined threshold.
The scheme is reported in Figure 1.

The matrix dimensions can be calculated taking in account the following considerations:

- the actual number of FFT boards installed inside the SIV is 6 and each one has 4×10^6 channels;
- each board can store the first 100 points that hit the threshold;
- every SIV file it is approximately 1 MB corresponding almost to 200 computed FFT

Notice that the dimension of the matrix is $M \in M(200, 24\times10^6)$, further we could consider it as a sparse matrix because the number of channels "ON" is really less than 24×10^6 (Figure 2).
We intend as a good ETI candidate every dopplered shifted monochromatic signal which turns out no to be a false alarm a radio interference (RFI), rejected by dedicated filters software subsequently described. From the literature that could be considered one of several possible forms of an intentional alien transmission.
In fact we would like to point out that, thanks to the open structure of the SALVE 2 program, a new pattern-recognition algorithm for different types of signals could be written and all the results merged at the end of the parallel flows.

§3 SALVE 2 Programs

The structure of the post processing software is shown in Figure 3. A scheduler controls and manages the entire process. *ACQUIRE* is the first application and it takes care of the data transfer from the PC Host to the base directory inside the local PC SERENDIP PROC1 disc. At the end of this phase, always under the control of the scheduler, the next modules start the interferences filtering and reducing with a particular dedicated software (*VFILTER*, *DFILTER*, *INTERFO*).

To get used to the local interference, we have developed a data display called *SCAN 6.6*. The study of the local RFIs has allowed us to develop the first version of algorithms for their treatment and rejection.

The *MERGE* file adds the possible candidates from the different pattern recognition algorithms (HOUGH, CORRELATION etc.) at the output level of the parallel flows already described.

All the alarms are stored day by day in the folder ALARM; an operator takes care of the analysis of the results.

To save space on disc, the SQUEEZE program "Zips" the binary data reducing the compressed file up to 60% with respect to the original file. It compresses all the SIV files.

§4 **From Radon to Generalized Hough Transform**

The Simple Hough Transform is a method used in literature to search for regular patterns [4,5], like lines in a two-dimensional set of data. For instance, every straight line in the Cartesian space (X,Y) is transformed by the Simple Hough transform in a peak (M,Q) in the bidimensional parameters space (Figure 4).

Thus, the difficult problem of looking for the straight line patterns in a plane is transformed by the Hough transform into the simpler problem of finding the local maximum in the two-dimensional space of parameters (M,Q). By accumulating the contributions of every couple of points belonging to the coherent by virtue of the Hough transform, it is possible look for the local maximum in the Hough space (M,Q) or in the equivalent (ρ,θ) polar space (Radon transform).

In mathematical terms, a straight line can be represented in the polar domain (ρ,θ) by means of the polar equation:

$$\rho = x\cos\vartheta + y\sin\vartheta$$

In this case, the Radon transform reads:

$$\mathrm{R}(\rho,\vartheta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{g(x, y)}\,\delta(\rho - \mathrm{x}\cos\vartheta - \mathrm{y}\sin\vartheta)\,\mathrm{dxdy}$$

where g(x,y) is the general function to transform, δ(x,y) is the Dirac δ function, ρ is the distance of the straight line from the origin of the axes and θ is the normal angle. This representation allows us to describe all the straight lines in the polar space within the range of $0 \le \vartheta < 2\pi$ e $\rho \ge 0$, as in Figure 4.

The Radon Transform is the simplest algorithm implemented for the actual postprocessing in SALVE2 programs; the initial Cartesian space (x,y) or "space of image" is the JTFM, then each point is transformed in the corresponding Radon Polar (RP) space R(ρ,θ).

In practice a real dopplered signal can be considered as a sinusoidal line in the Cartesian space (Figure 5); a straight line can be approximated only for short time periods.

In fact, supposing the JTFM to be compensated with respect to its own reference frame, the hypothetical alien transmitter, in comparison to the system geometric earthling, would only show the motions of rotation and revolution around its star.

The second version of the algorithm (Generalized Hough Transform), that is a generalization of the previous one, differs from it for the simple but expensive fact which, instead of seeking straight lines in the space image, sinusoids or sinusoidal arcs are sought. In general that can be of the type:

$$f(x) = A \cdot \sin(P \cdot x - F) - q$$

where (A, P, F, q) are the main vectors of Generalized Hough (GH) space and correspond to:

A: amplitude of sinusoid

P: angular velocity $\frac{2p}{t}$ where τ is the period corresponding to the Doppler shift of the signal.

F: Initial Phase

q: Axis Offset of sinusoid

As is it possible to understand, the newest GH space has four dimensions, increasing the computational complexity with respect to the standard Radon algorithm.

To physically avoid impossible solutions and therefore to decrease the time of calculation and the engaged memory, a subset of the quadridimensional dominion is chosen (the express integer in table here under has been considered as the number of columns or lines of the JTFM):

Parameter	*Range*
A amplitude	[0, 1000]
P angular velocity	$\frac{2p}{t}$ where $\tau \in [35\text{min}, 14.7\text{h}] \rightarrow$ [1250, 31250]
F phase	[0, π]
q offset	[-1000, 1000]

The transformation could be considered as:

$$\forall(A, P, F, q)$$

$$H(A, P, F, q) = \int_{-\infty}^{+\infty} g(x, A\cdot\sin(P\cdot x - F) - q)\cdot dx =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \text{g(x, y)}d\,(y - A\cdot\sin(P\cdot x - F) - q)\ \text{dxdy}$$

Such function is not null when the (x,y) fits the function $y - A\cdot\sin(P\cdot x - F) - q$ or when, once fixed the parameters (A,P,F,q) , the couple (x,y) is a solution of the equation:

$$y - A\cdot\sin(P\cdot x - F) - q = 0$$

that in the space of the image represents a sinusoid while, in the GH space, is an hypersurface of dimension three. The intersection, for instance, of two of these hypersurfaces behaves as a double value in the points of intersection, in comparison to those that don't belong to this set, and therefore it is here that one should look for the possible candidates to be the desired sinusoids. At this point it is possible to pass in rev

iew the GH space and, with appropriate filters and algorithms, to find out the global maximum.

To improve the search of the maximum and to avoid that the discretization of the space image could introduce false alarms, a clusterization K-MEANS is used. Defining at least a radius of tolerance of the hyperspheres of the transformed points, all the transformed points of the space image whose hyperspheres intersect it can be considered as a single cluster.

To better define the parameters (A, P, F, q) of the cluster, an interactive process can define a common barycenter and associate an average vote to it.

§5 **Simulations**

Due to the complexity of the described algorithm, some tests have been made on a simulated matrix.

The sinusoids shown in the next Figure 6, Figure 7 are the maximums found in the GH space and again represented in the image space in order to understand if the algorithm really found the desired function.

§6 CONCLUSIONS

Within the SETI project a new algorithm is implemented for the sinusoidal pattern recognition of signals in the Join Time Frequency Matrix.
The used mathematical tool is the four dimensional Generalized Hough transform.
Some tests have been made with some simulated data. The results were extremely satisfactory despite the dimension of the space image was smaller because of the computing power limits.
With the due adaptations to more powerful computers, the algorithm will be implemented in Italian SALVE2 programs and, in a near future, it could be used for the most advanced SETI searches.

References

[1] S. Montebugnoli et al. "SETItalia" A new Era in Bioastronomy ASP Conference Series, Vol 213, 2000, pages 501-504

[2] J. Monari et al. "SALVE 2 (Software Aimed at off Line Verification ETI signals ver 2.0)" International Astronatuical Proceedings IAA-01-IAA.9.1.10.

[3] D. Werthimer et al: "The Serendip IV Arecibo Sky Survey" A new Era in Bioastronomy ASP Conference Series Vol 213, 2000, pages 479-483.

[4] Clark F. Olson "An implementation of Hough transform for curves detections", Computer vision and Image Understanding, Vol 73, 1999.

[5] F. Evans et al. " A survey comparison of Hough transform", Proc IEEE Comp. Soc. Workshop Computer Architecture for Pattern Analysis and Image Database Management, 1985, pages 278-380.

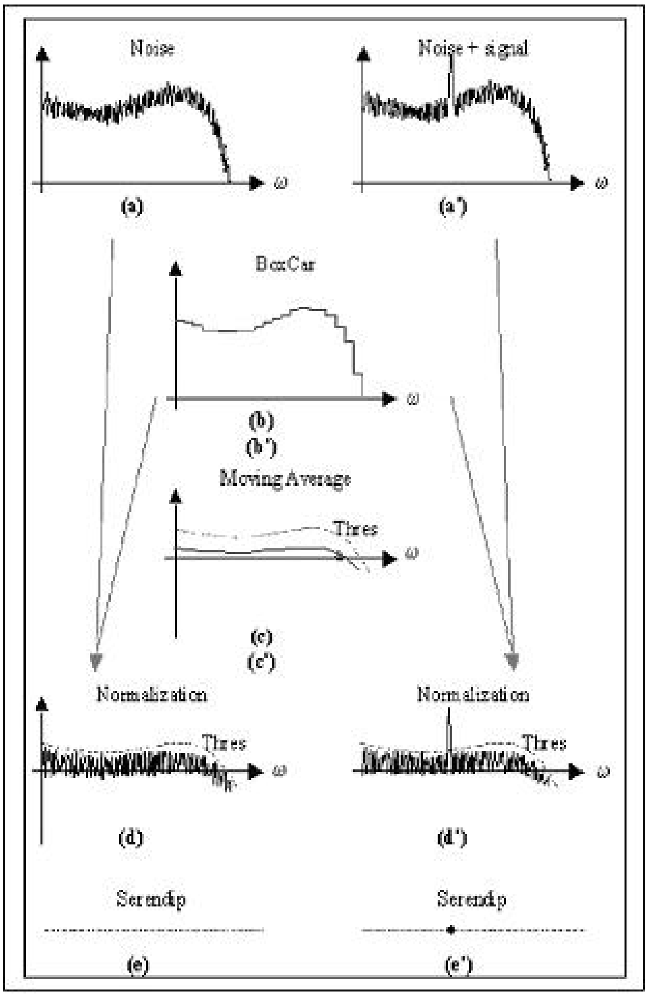

Figure 1: JTF Analysis; each channel is "ON" only when a signal hits a particular threshold set by user

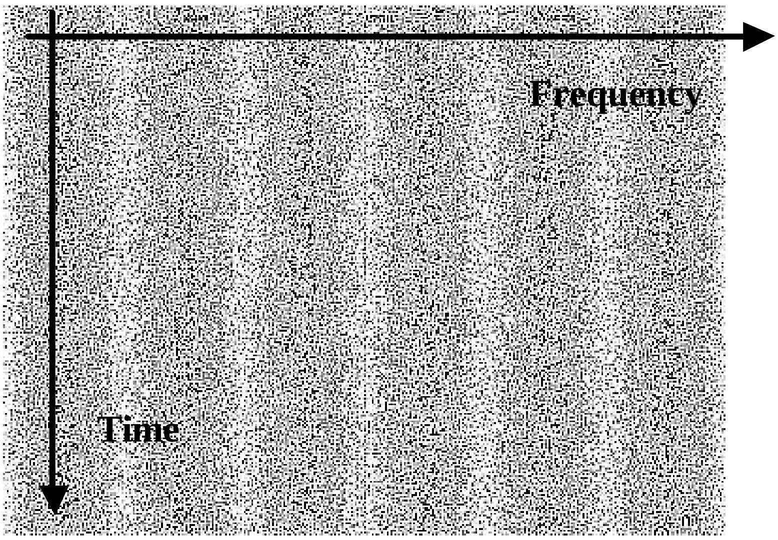

Figure 2: JTFM; the dimension of the matrix is $M \in M(200, 24 \times 10^6)$,

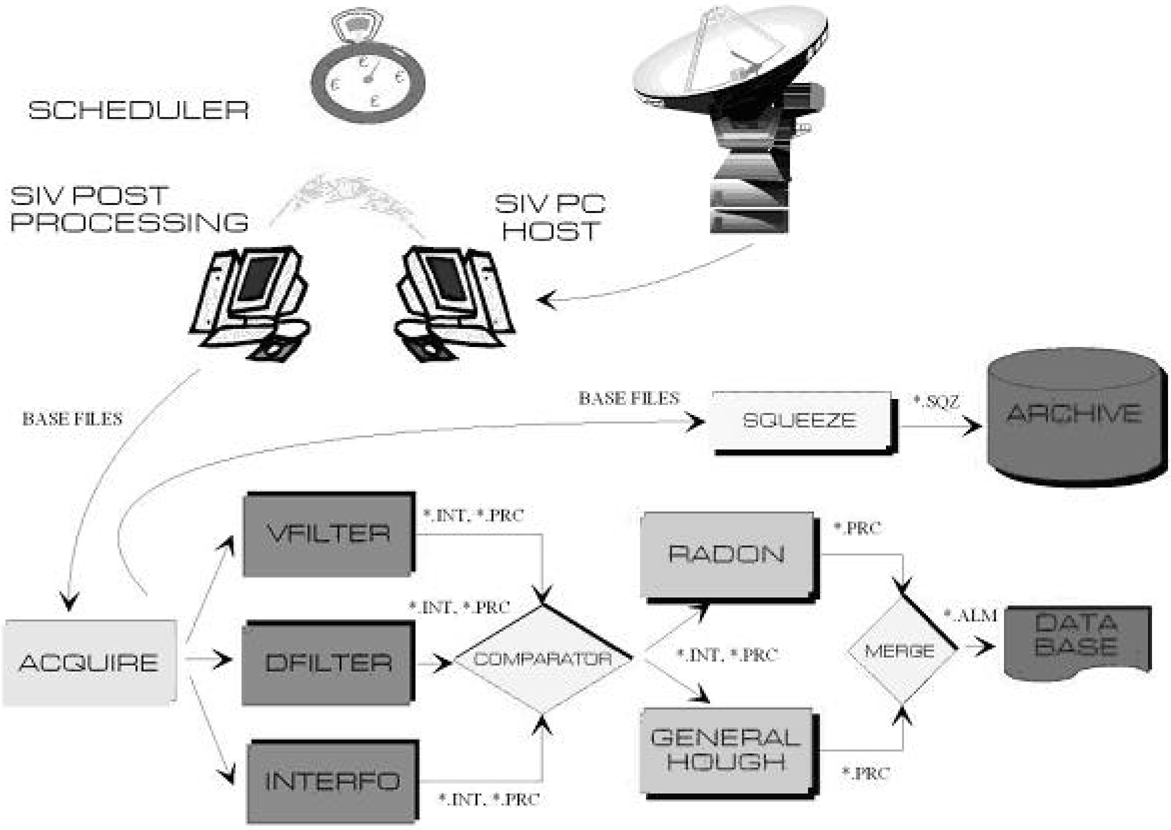

Figure 3: SALVE2 Programs; structure of the post processing software

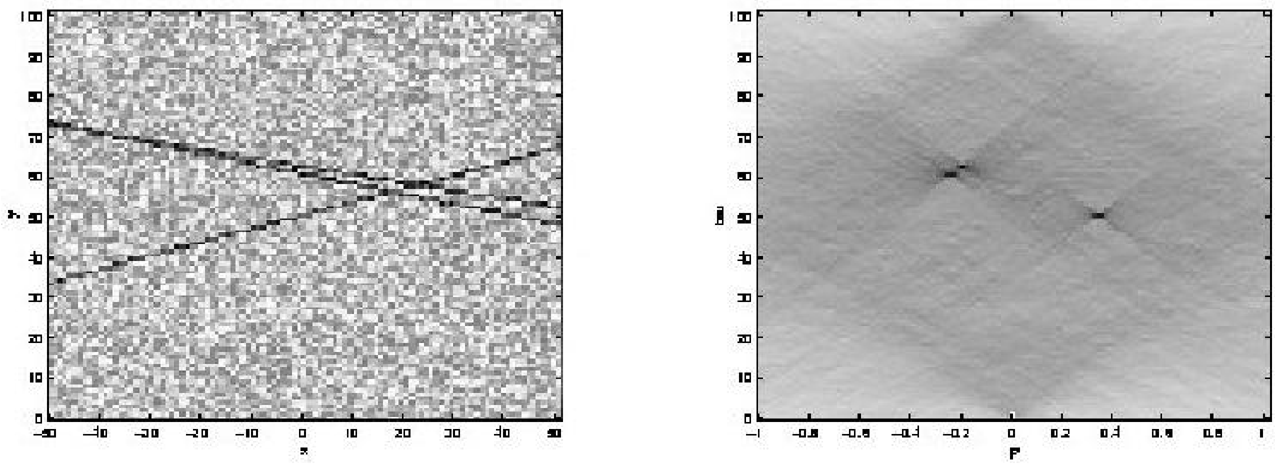

Figure 4: Simple Radon transform.

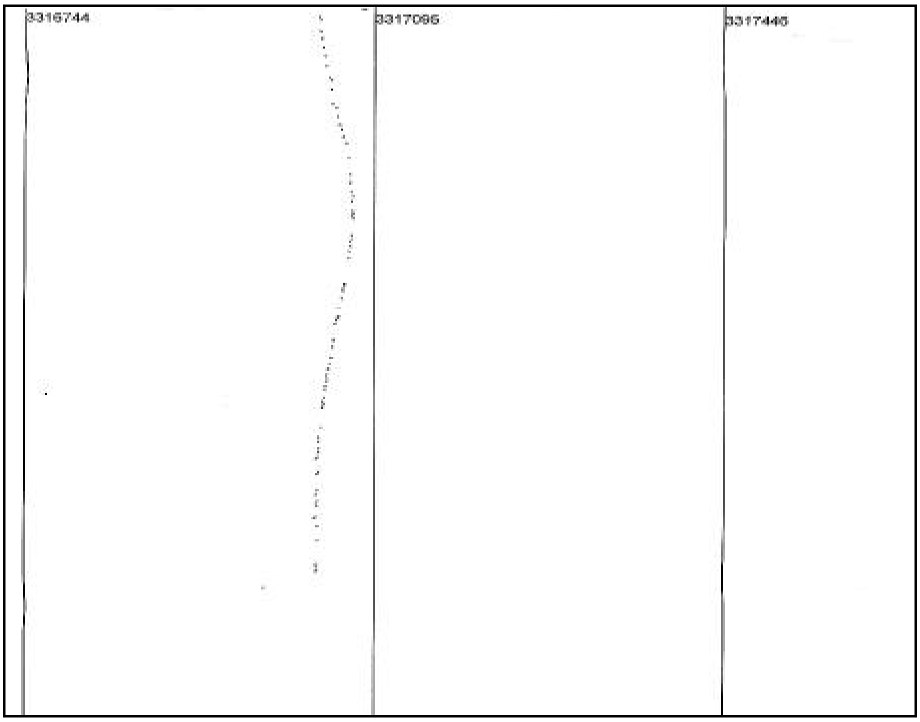

Figure 5: Is that a WOW signal received from Italian radiotelescope (7 July 1998)?

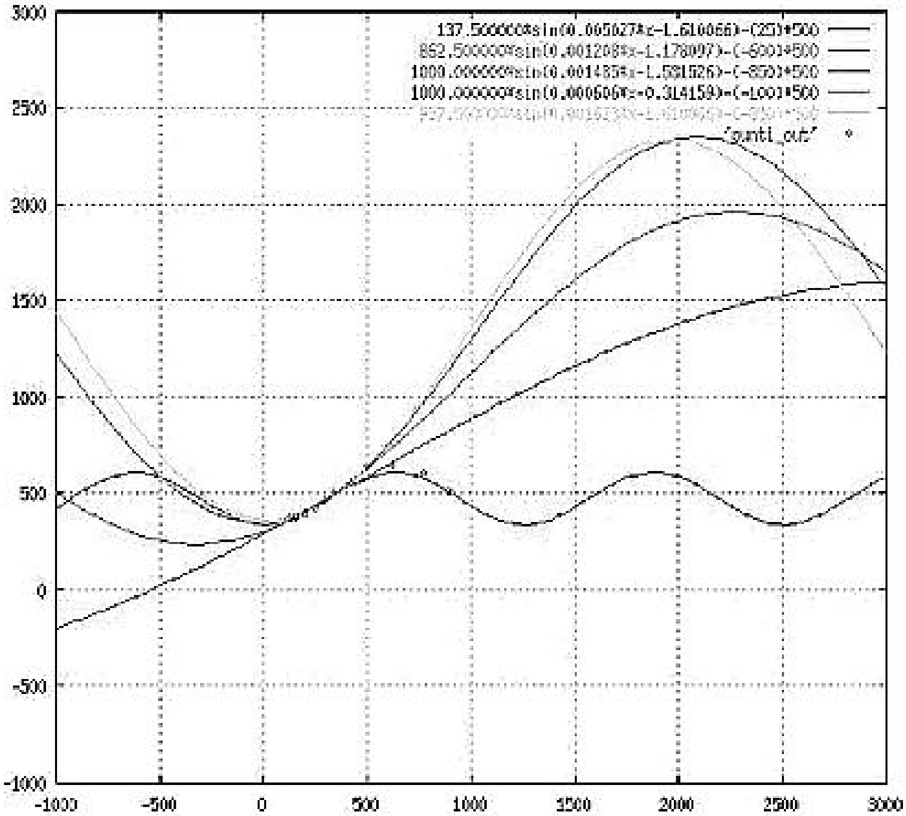

Figure 6: Solutions found by GHT algorithm. The most probable is the first one that it perfectly matches with the simulated sinusoid set by the user.

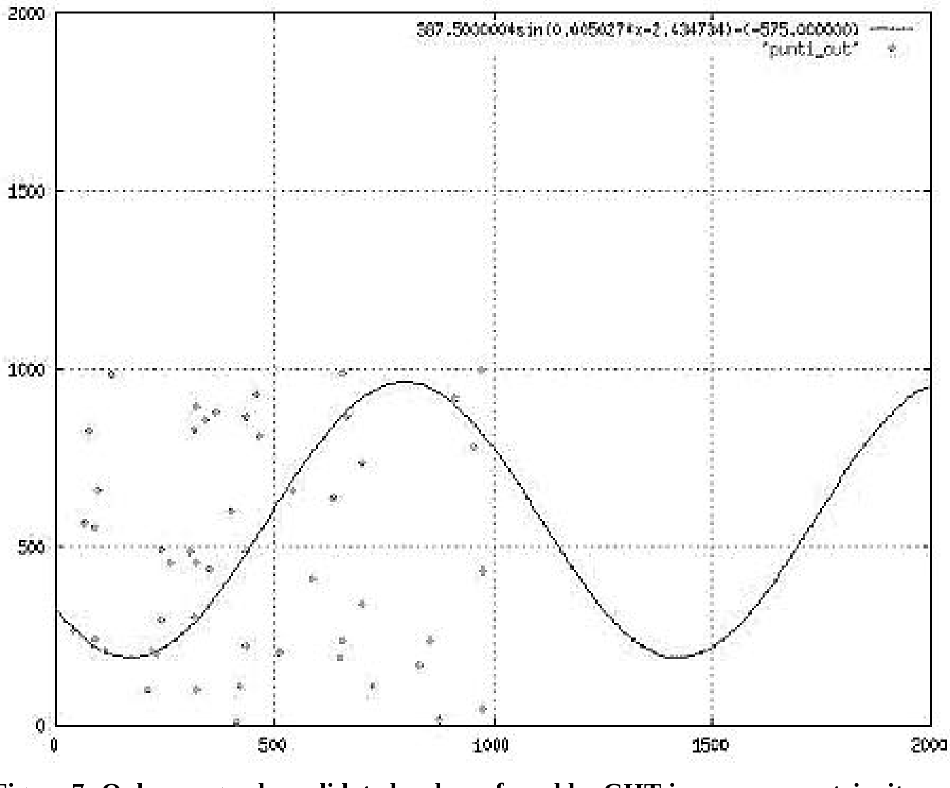

Figure 7: Only one good candidate has been found by GHT in a sparse matrix, it perfectly matches with the simulated sinusoid set by the user.