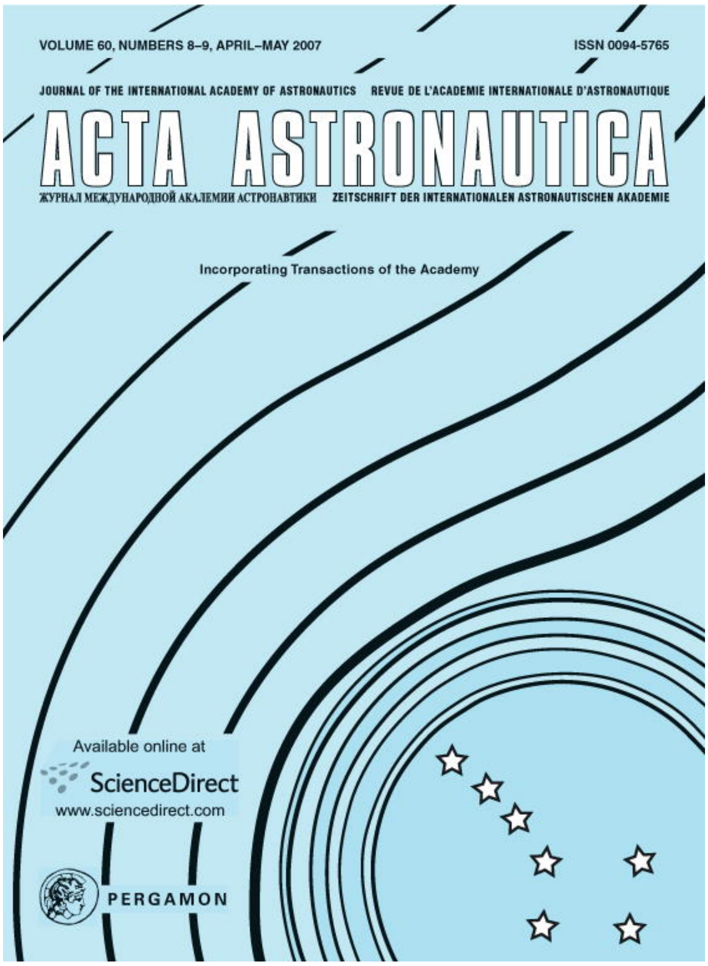
VOLUME 60, NUMBERS 8–9, APRIL–MAY 2007
ISSN 0094-5765
JOURNAL OF THE INTERNATIONAL ACADEMY OF ASTRONAUTICS
REVUE DE L'ACADEMIE INTERNATIONALE D'ASTRONAUTIQUE
ACTA ASTRONAUTICA
ЖУРНАЛ МЕЖДУНАРОДНОЙ АКАДЕМИИ АСТРОНАВТИКИ
ZEITSCHRIFT DER INTERNATIONALEN ASTRONAUTISCHEN AKADEMIE
Incorporating Transactions of the Academy
Available online at
ScienceDirect
www.sciencedirect.com
PERGAMON

PERGAMON

Available online at www.sciencedirect.com
ScienceDirect

Acta Astronautica 60 (2007) 775–779

ACTA ASTRONAUTICA

www.elsevier.com/locate/actaastro

Academy transactions note

Advantages of Karhunen–Loève transform over fast Fourier transform for planetary radar and space debris detection

Claudio Maccone*

International Academy of Astronautics, Via Martorelli 43, Torino (TO) 10155, Italy

Available online 27 October 2006

Abstract

The present article describes that the range of any radiotelescope (and radar in general) may be increased by virtue of software, if one replaces the fast Fourier transform by the Karhunen–Loève transform. The range increases with the inverse of the fourth root of the signal-to-noise ratio when this ratio decreases. Thus, the range on any radiotelescope (and radar) may be increased without changing the hardware at all, but by changing the software only. This improvement in the range of the radiotelescope is currently implemented at the 32-m antenna located at Medicina, near Bologna, in Italy, for both SETI and general radioastronomy.

1. Introduction: two examples of bistatic planetary radars

Since the 1990s, several planetary radar experiments have been carried out all over the world (for instance, see Refs. [1–4]). In order to further improve these systems, a somewhat innovative idea is proposed in this paper: increasing the distance where the planetary radar can still reach (i.e. the range) by virtue of a pure software trick: the replacement of the good old fast Fourier transform (FFT) by virtue of the more recent and mathematically superior Karhunen–Loève transform (KLT) [5]. What the KLT is, in mathematical terms, will be briefly described in Sections 3 and 4 of this paper. At the moment, we start by showing that, upon keeping the whole hardware just the same, the range r of any antenna increases inversely to the fourth root of the decreasing signal-to-noise ratio (SNR). This result follows at once from the equation yielding the cross section $\sigma_{\min}$ that can be detected by any bistatic radar (for a proof of this equation, see, for instance, the web site: http://wps.cfc.forces.gc.ca/papers/csc/csc29/mds/hudson.htm)

$$\sigma_{\min} = \frac{4\pi r_1^2 r_2^2 \lambda^2 k T_{\text{sys}} \text{SNR}}{P_{\text{t}} S_{\text{t}} S_{\text{r}} t}. \tag{1.1}$$

In this equation one has $\sigma_{\min}$ the minimal detectable radar cross section; r_1 the distance between the transmitting antenna and the target in space (i.e. space debris or even a small asteroid passing nearby); r_2 the distance between the receiving antenna and the target in space; λ the wavelength of the radio beam emitted by the transmitting radiotelescope (or radar); k the Boltzmann constant $=1.380658 \times 10^{-23}$ J/K; T_{sys} the system temperature of the receiving antenna; SNR the signal-to-noise ratio at the receiving antenna; P_{t} the power of the transmitting antenna; S_{t} the effective area of the transmitting antenna; S_{r} the effective area of the receiving antenna; and t the integration time at the receiving antenna.

As a practical, numerical example of Eq. (1.1), let us consider the important case when the small object that

* Tel.: +39 011 2055 387; mobile: +39 3471053812.
E-mail address: clmaccon@libero.it.
URL: http://www.maccone.com/

0094-5765/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.actaastro.2006.08.015

we want to detect is located in geostationary Earth orbit (GEO), namely

$$r_1 = 35{,}786\,\text{km} = r_{\text{GEO}},\qquad r_2 = 35{,}786\,\text{km} = r_{\text{GEO}}. \tag{1.2}$$

Next, suppose for a moment that the remaining numbers in Eq. (1.1) are those that applied in 2001 to the first Italian planetary radar experiment described in Ref. [4]. Then, the transmitting antenna was the 70-m dish located at Evpatoria in Crimea (now part of the Ukraine), with typical data (in C-band, in which the transmission occurred):

$$\begin{aligned} S_{\text{t}} &= S_{\text{Evpatoria}} = 2520\,\text{m}^2,\\ P_{\text{t}} &= P_{\text{Evpatoria}} = 150\,\text{kW},\\ \nu_{\text{Evpatoria}} &= 5010.024\,\text{MHz},\\ \lambda_{\text{Evpatoria}} &= 5.984\,\text{cm}. \end{aligned} \tag{1.3}$$

The receiving antenna was the 32-m dish located at Medicina (near Bologna, in Italy), with typical data (notice that the dish effective area changes according to the frequency, C-band in this case):

$$\begin{aligned} S_{\text{r}} &= S_{\text{Medicina\_in\_C\_band}} = 466\,\text{m}^2,\\ T_{\text{sys}} &= T_{\text{Medicina}} = 50\,\text{K},\\ t_{\text{Medicina}} &= 10\,\text{s},\\ \text{SNR}_{\text{Medicina}} &= 10. \end{aligned} \tag{1.4}$$

Then, by replacing Eqs. (1.2)–(1.4) into Eq. (1.1), one gets the minimal cross section $\sigma_{\min}$ that would be detectable in GEO by virtue of the Evpatoria–Medicina bistatic radar

$$\sigma_{\min} = 2.892 \times 10^{-4}\,\text{m}^2 \tag{1.5}$$

corresponding to an object having the size

$$d_{\min} = \sqrt{\frac{4\sigma_{\min}}{\pi}} = 1.919\,\text{cm} \approx 2\,\text{cm}. \tag{1.6}$$

In other words, the Evpatoria–Medicina bistatic radar would be capable of detecting space debris objects like a bolt ($\sim 2\,\text{cm}$) in GEO.

Even smaller objects one would detect in GEO by virtue of a Goldstone–Medicina bistatic radar. In fact, the Goldstone antenna in the Mojave desert (California, USA), has the typical data for a transmission in X-band:

$$\begin{aligned} S_{\text{t}} &= S_{\text{Goldstone}} = 2694\,\text{m}^2,\\ P_{\text{t}} &= P_{\text{Goldstone}} = 460\,\text{kW},\\ \nu_{\text{Goldstone}} &= 8560.0\,\text{MHz},\\ \lambda_{\text{Goldstone}} &= 3.502\,\text{cm}. \end{aligned} \tag{1.7}$$

The Medicina data in this case are slightly different, for a smaller effective area applies to the X-band receiver:

$$\begin{aligned} S_{\text{r}} &= S_{\text{Medicina\_in\_X\_band}} = 389.3\,\text{m}^2,\\ T_{\text{sys}} &= T_{\text{Medicina}} = 50\,\text{K},\\ t_{\text{Medicina}} &= 10\,\text{s},\\ \text{SNR}_{\text{Medicina}} &= 10. \end{aligned} \tag{1.8}$$

The conclusion is that the smallest detectable object in GEO by virtue of the Goldstone–Medicina bistatic radar would have a size smaller than 1 cm, namely

$$d_{\min} = \sqrt{\frac{4\sigma_{\min}}{\pi}} = 0.679\,\text{cm} \approx 0.7\,\text{cm}. \tag{1.9}$$

2. Range vs. SNR in radiotelescopes and radars

We can now propose the leading idea of this paper: the range of any radiotelescope and/or radar may be increased by changing the software only (that is, without changing anything in the hardware at all) if one replaces the good old FFT by virtue of the newer KLT.

Consider Eq. (1.1) again. Upon replacing

$$r_1 = r_2 = r \tag{2.1}$$

and then solving for r, one gets immediately the expression of the range r as a function of the SNR at the receiving station:

$$r(\text{SNR}) = \sqrt[4]{\frac{\sigma_{\min} P_{\text{t}} S_{\text{t}} S_{\text{r}} t}{4\pi\lambda^2 k T_{\text{sys}}}}\,\frac{1}{\sqrt[4]{\text{SNR}}}. \tag{2.2}$$

In other words, upon keeping everything else unchanged (the hardware), the range increases as the inverse of the root to order four (i.e. the fourth root) of the SNR at the receiving station.

The KLT does just that: it lowers the SNR of the feeblest detectable signals, when replaced in the software instead of the FFT.

Fig. 1 shows the increase in the range of the Evpatoria–Medicina planetary radar (as described in Section 1) when the KLT is used instead of the FFT in the filtering software at Medicina, which is what the Medicina Radiotelescope Team (of which this author is an external co-worker) has been doing since the late

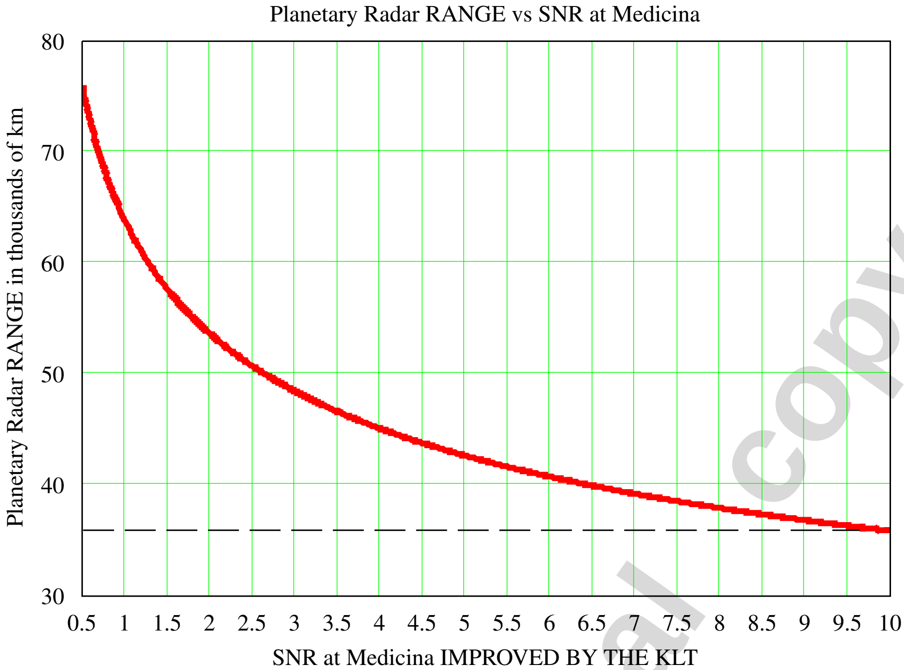

Fig. 1. Range vs. SNR for the planetary radar at Medicina, near Bologna, in Italy. The improvement in the range is due to the replacement of the FFT by virtue of the KLT (Karhunen–Loève transform) as the noise filtering tool. The horizontal dashed line is the GEO distance of 35,786 km above the Earth. The solid curve increases like the inverse of the fourth root of the SNR as long as the SNR decreases from 10 to 0.5 because of the superior filtering provided by the KLT over the FFT.

1990s under the direction of Ing. Stelio Montebugnoli (see, for instance, Ref. [8]).

But what is the KLT?

The KLT is a mathematical tool superior to the FFT in that it rigorously applies to any finite bandwidth, rather than applying to infinitely small bandwidths only (i.e. to monochromatic signals) as the FFT does. Also, the KLT applies to both stationary and non-stationary processes, and when the background noise distribution is colored, rather than just white.

The KLT is described more in detail (but without a full, appropriate mathematical treatment, that would take too many equations) in Sections 3 and 4 of this paper.

3. A heuristic introduction to the KLT

The KLT is a rather recent mathematical tool capable of improving our understanding of physical phenomena, and it is superior to the classical FFT, as intuitively described by the following mechanical analogy.

Consider an object, for instance a book, and a three-axes rectangular reference frame, oriented in an arbitrary fashion with respect to the book. Then, Newtonian mechanics shows that all mechanical properties of the book are described by a 3×3 symmetric matrix called the "inertia matrix" (or "inertia tensor") whose elements are, in general, all different from zero. Now, handling a matrix whose elements are all non-zero, is obviously more complicated than handling a matrix where all entries are zeros except for those on the main diagonal (i.e. a "diagonal matrix"). Thus, one may be led to wonder whether a certain transformation of axes exists that changes the inertia matrix of the book into a diagonal matrix. Newtonian mechanics shows then that only one such privileged orientation of the frame with respect to the book exists yielding a diagonal inertia matrix: the three axes must coincide with a set of three axes (parallel to the book edges) called "principal axes" of the book, or "eigenvectors" or "proper vectors" of the inertia matrix of the book. In other words, each body possesses an intrinsic set of three rectangular axes that describes its dynamics at best. And one can always compute the position of the eigenvectors with respect to a generic reference frame by means of a certain mathematical procedure called "finding the eigenvectors of a square matrix".

Now let us go over to signal processing, which is the main theme of this paper. By adding random noise to a deterministic signal one obtains what is called a "noisy signal" or, in case the power of the signal is much less than the power of the noise "a signal buried into the noise". Since the signal + noise is a random function of the time, denoted hereafter by $X(t)$, one can describe it well by a statistical quantity called autocorrelation (or simply correlation), defined as the mean value of the product of the values of $X(t)$ at two different instants t_1 and t_2 and formally written $E\{X(t_1)X(t_2)\} \equiv \langle X(t_1)X(t_2)\rangle$. This correlation, obviously symmetric in t_1 and t_2, may play just the same role as the inertia matrix in the book example. Thus, if one firstly seeks for the eigenvectors of the correlation, and then changes the reference frame over to this new set of vectors, the simplest possible description of the signal + noise is achieved.

The next step is the rearranging of the eigenvalues in decreasing order of magnitude and, consequently, also the rearranging of the eigenvectors corresponding to each eigenvalue. There is no degeneracy, i.e. only one eigenvector corresponds to each eigenvalue. Also, all eigenvalues turn out to be positive, and so, once rearranged, they form a decreasing sequence whose first eigenvalue is the largest one, called the "dominant" eigenvalue by mathematicians.

We are now ready to compute the direct KLT transform of the signal + noise: simply use the new set of eigenvectors to describe the signal + noise: the signal + noise in the new representation is just the direct KLT transform of the older signal + noise. So, the KLT is just a linear transformation of axes: nothing easier than that! But what is its statistical meaning? Well, since the eigenvalues also are the variances of the zero-mean set of data, this means that we are ordering the axes according to their decreasing order of statistical importance. In other words, the first eigen-axis is the one around which the variance is largest. The second eigen-axis, is the one with second largest variance, and so on. In other words still, the more eigenvectors one takes into account, the more one "grabs" out of the statistically significant part of the data. But, since the variances around the axes decrease as long as one takes into account more and more axes, one is really "grabbing" less and less statistically significant stuff. This "feeling" is the key to the KLT filtering.

In fact, the KLT filtering simply consists in only taking a small, finite number of eigenvectors out of the set of all (infinite) eigenvectors, and then declaring the part of the data spanned by this smaller set of eigenvectors as the "statistical bulk", or the "signal", out of the original signal + noise. The "noise" is then automatically the cut-away part. Finally, in order to recover the signal out of the noise, one has simply to back-transform, or inverse KLT, the small set of data that has been regarded as the statistically "significant" part of the original, full signal + noise.

And if the input is an image, rather than a noisy signal, then the KLT bulk is just the KLT-compressed image.

So, the KLT may be equally well used for both noise filtering and data compression.

4. The KLT expansion

The KLT is named for two mathematicians, the Finn Kari Karhunen and the French–American Michel Loève (1907–1979), who proved, at about the same time (1947) and independently, that the series (4.1) hereafter is convergent. When put this way, the KLT looks like a purely mathematical topic, but this is not, of course, the case. Instead, we are going to use the language familiar to engineers and radioastronomers, and so we shall say that it is possible to represent the signal + noise $X(t)$ as the infinite series (called KLT expansion)

$$X(t) = \sum_{n=1}^{\infty} Z_n \phi_n(t), \tag{4.1}$$

where the Z_n are just random variables (i.e. they are not stochastic processes) and the $\phi_n(t)$ are just ordinary (i.e. deterministic) time functions. Assuming that the signal + noise autocorrelation $E\{X(t_1)X(t_2)\} \equiv \langle X(t_1)X(t_2)\rangle$ is a known function of t_1 and t_2, it can be proven that the functions $\phi_n(t)$ $(n = 1, 2, \ldots)$ are the eigenfunctions of the correlation. In other words, the correlation is treated as an operator acting on the time variable, and its eigenfunctions are the solutions to the integral equation

$$\int_0^T \langle X(t_1)X(t_2)\rangle \phi_n(t_1)\,\mathrm{d}t_1 = \lambda_n \phi_n(t_2). \tag{4.2}$$

These $\phi_n(t)$ form an orthonormal basis in the Hilbert space, and they actually are the best possible basis to describe the signal + noise, better than any classical Fourier basis made just of sines and cosines. One can thus say that the KLT adapts itself to the shape of the signal + noise, whatever it is. A further advantage of KLT is that the Z_n in Eq. (4.1) are orthogonal random variables in the statistical sense, i.e. $\langle Z_m Z_n\rangle = \delta_{mn}\lambda_n$ i.e. they are uncorrelated. In addition, if $X(t)$ is Gaussian, this amounts to statistical independence, i.e. the terms in the KLT expansion are statistically independent of each other. Finally, since the constants λ_n are both the (positive) eigenvalues and the variances of the random variables Z_n, the KLT expansion, when truncated to keep only the first few terms, may be proven to be the best approximation to the full KLT expansion in the mean square sense.

A much more detailed analytical treatment of the KLT for general, *non-stationary* stochastic processes may be found in the author's book [5], embodying also the relativistic KLT, i.e. the KLT of objects in relativistic motion with respect to Earth, such as highly red-shifted quasars or even future relativistic starships. Good references about the KLT of *stationary* stochastic processes only, such as the inputs to radiotelescopes are assumed to be in SETI, are in Refs. [6–8].

5. Conclusion

We describe here that the range on any radiotelescope and/or radar increases with the inverse of the fourth root of the SNR when the SNR is reduced.

The SNR is thus reduced without touching the hardware at all, simply by replacing the FFT by virtue of the KLT.

Although the computational burden of the KLT is higher than the FFT, the great contemporary advances in computer technology now enable the adoption of the KLT at moderate costs and with the above-mentioned advantages.

References

[1] S.J. Ostro, et al., Radar observations of Asteroid 1620 Geographos, Icarus 121 (1996) 46–66.
[2] R.S. Hudson, et al., High-resolution model of Asteroid 4179 Toutatis, Icarus 161 (2003) 346–355.
[3] S.J. Ostro, et al., Radar detection of Asteroid 2002 AA29, Icarus 166 (2003) 271–275.
[4] M. Di Martino, et al., Results of the first Italian planetary radar experiment, Planetary and Space Science 52 (2004) 325–330.
[5] C. Maccone, Telecommunications, KLT and Relativity, vol. 1, IPI Press, Colorado Springs, CO, USA, 1994 (ISBN 1-880930-04-8).
[6] F. Biraud, SETI at the Nançay Radiotelescope, Acta Astronautica 10 (1983) 759–760.
[7] R.S. Dixon, On the detection of unknown signals, in: S. Shostak (Ed.), Third Decennial US–USSR Conference on SETI, ASP Conference Series, vol. 47, 1993, pp. 129–140.
[8] S. Montebugnoli, SETI-Italia 2003: status review, Paper IAC-03-IAA.9.1.02 presented at the International Astronautical Congress (IAC) held in Bremen, Germany, September 29–October 3, 2003.