10.1515/astro-2017-0165

Baltic Astronomy, vol. 22, 347–361, 2013

COMPRESSIVE SENSING: ANALYSIS OF SIGNALS IN RADIO ASTRONOMY

G. Gaigals[1], M. Greitāns[2] and A. Andziulis[3]

[1] *Engineering Research Institute "Ventspils International Radio Astronomy Centre", Ventspils University College, Inženieru iela 101, Ventspils, LV-3601, Latvia; gatisg@venta.lv*

[2] *Institute of Electronics and Computer Science, 14 Dzerbenes St., Riga, LV-1006, Latvia; modris_greitans@edi.lv*

[3] *Klaipėda University, Manto 84, Klaipėda, LT-92294, Lithuania; llii.063.jrtc@gmail.com*

Received: 2013 September 19; accepted: 2013 November 5

Abstract. The compressive sensing (CS) theory says that for some kind of signals there is no need to keep or transfer all the data acquired accordingly to the Nyquist criterion. In this work we investigate if the CS approach is applicable for recording and analysis of radio astronomy (RA) signals. Since CS methods are applicable for the signals with sparse (and compressible) representations, the compressibility of RA signals is verified. As a result, we identify which RA signals can be processed using CS, find the parameters which can improve or degrade CS application to RA results, describe the optimum way how to perform signal filtering in CS applications. Also, a range of virtual LabVIEW instruments are created for the signal analysis with the CS theory.

Key words: methods: radio astronomy signals: compressive sensing, sparsity, filtering

1. INTRODUCTION

For the research of distant and faint radio astronomical (RA) objects, about two tens of radio telescopes are involved in observing sessions in the Very Long Basis Interferometry (VLBI) project. In a single session an individual station usually records about ten GB of observational data. Such big amount of data is received because of high sampling rate which is required by the Nyquist criterion. In ordinary VLBI sessions, hard-drive data disks are sent to the correlator station for processing and imaging of radio sources. In the case of modern e-VLBI, there is no need for post office utilities, but there is need for a good network infrastructure to handle data transfer without interruptions. In the case of creation of an observational archive there is also need for a sufficient storage space.

The compressive sensing (CS) today is a quite attractive technique as the search in Springer[1] gives more than 12 thousand hits. The CS-based methods are used in many applications and research areas, for example, image denoising, artificial

[1] http://link.springer.com/search/query=compressive+sensing

intelligence, machine learning. By its nature CS is used in RA as an effective tool for image synthesis (deconvolution) and image quality improvement (denoising) (Li et al. 2011; Suksmono 2009; Högbom 1974; Wiaux et al. 2009; Wenger et al. 2010).

As is shown in Donoho (2006), Candès et al. (2006a) or Candès & Romberg (2007), the application of CS is determined by the signal sparsity or compressibility – if the signal in some representation is sparse, then it is compressible and vice versa. Although in Springer about three hundred articles are related to CS and RA, no articles are devoted to compressive sampling of RA signals. The aim of this work is a research of compressibility of RA signals to find if CS methods are applicable to RA signals directly.

We start from the analysis of RA signal classes to identify signal types for which CS methods can be successfully applied. Then the necessary instruments for RA signal analysis in LabVIEW are developed, in particular a tool for the analysis of signal entropy, a tool for the analysis of sparsity of signal transformation spectra, and others. Different RA signals for the analysis were collected by taking previous RA observational data and by new observations of the selected space objects.

The most important results are: nearly all RA signals are compressible, only some specific RA signals are sparse, the signal spectrum cannot represent its sparsity, the signal filtering in CS can be done without a dedicated signal filtering stage in the signal path. The analysis of CS measurement matrix and specific steps in the signal recovery can help to improve the quality of the signal.

2. THEORY

2.1. Compressive sensing

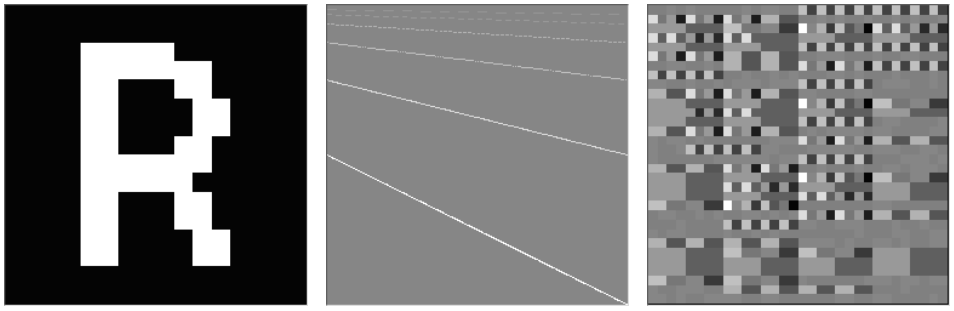
In overall CS can be described using a citation from Fornasier & Rauhut (2011): "Compressive sensing is a new type of sampling theory, which predicts that sparse signals and images can be reconstructed from what was previously believed to be incomplete information. As a main feature, efficient algorithms such as $\ell 1$-minimization can be used for recovery." In the paper only most important definitions and statements are given to show why such methodology is chosen. The complete theory is described in Candès et al. (2006a), Candès & Romberg (2007) and Fornasier & Rauhut (2011).

2.1.1. Sparse signals

For a given signal x there can be a lot of transforms (representations), where the transform coefficient vector c has only a small amount k of non-zero entries. In the matrix form:

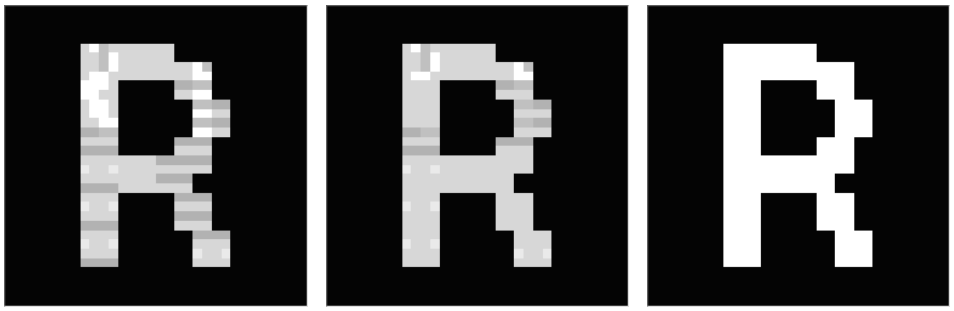
$$x = \Psi\ c, \tag{1}$$

where Ψ is so-called transform or sparse basis matrix with the dimensions $N \times N$, where N is the total number of samples in x. Figure 1 shows the transformation of a picture of letter "R" (honouring experiments at the Rice University) with the Haar matrix. Figure 2 shows the reconstruction results using Eq. (1), where different amount of non-zero coefficients are taken for recovery (coefficients are thresholded by a certain value for k limiting). The peak signal-to-noise ratio (PSNR) was calculated to show the degree of degradation of the reconstructed image.



(a) (b) (c)

Fig. 1. Transformation of a signal using the Haar basis: (a) binary picture of "R" with dimensions 32×32 pixels; (b) the Haar matrix; (c) the picture representation in the transformed domain.



(a) (b) (c)

Fig. 2. The restored picture after thresholding of the transform coefficients (see Figure 1c): (a) $k = 558$, PSNR = 30,32 dB; (b) $k = 592$, PSNR = 34,64 dB; (c) $k = 632$, perfect recovery.

As can be seen in the Haar representation (Figure 1) there is a lot of non-zero entries in c; k is not small so in this case the Haar transform (HT) cannot be called as a sparse representation of x. At the same time, the Walsh transform (WT), which uses the ordered Walsh matrix, can be called as a sparse representation of x because just for $k = 256$ it gives a perfect recovery (Figure 3a). The discrete cosine transform (DCT) does not give a perfect recovery in this case even for k close to N because of a binary nature of the original signal (Figure 3b).

The process of applying the transform to a signal is called sparsifying transform (Lustig et al. 2007). Actually, for known x and Ψ this means solving of (1) for c if the transform coefficients are needed. The number of non-zero entries k in c does represent the signal sparsity – it is said that c is k-sparse, but the signal x can be also called as k-sparse because it has a k-sparse representation.

2.1.2. Identity transform

RA signals usually are very weak, in Very Long Basis Interferometry (VLBI) their level usually is below the noise level of receiving system and the background noise, so usually RA signals are recorded only with a couple of quantization levels.

(a) (b)

Fig. 3. The restored picture after thresholding with different transforms: (a) WT, $k = 256$, perfect recovery; (b) DCT, $k = 724$, PSNR = 57.24 dB (artifacts in the picture are enhanced).

Signal is captured in a way that keeps average value of the signal close to 0.5, where 1 is the logical level '1' of the capture hardware. Such signal can be treated as sparse because in the signal usually a half of the readings are zero. Similar assumptions are used in medical image processing (Lustig et al. 2997). To use CS methods original signal can be "transformed" to sparse representation using the identity transform by taking in (1) $\Psi = I$, where I is the identity matrix.

2.1.3. Sensing of the signal

The signal sensing is being done by calculating so-called measurement vector:

$$y = \Phi\ x, \tag{2}$$

where Φ is the measurement matrix with dimensions $M*N$, where M is determined by properties of x, Φ and Ψ; usually $M \ll N$ (Donoho 2006; Candès et al. 2006a; Candès & Romberg 2007; Fornasier & Rauhut 2011; Lustig et al. 2007). The values of Φ rows are the measurement basis vectors ϕ_m $(m = 1..M)$, while the values of the Ψ columns are the transformation or sparse basis vectors ψ_n $(n = 1, ..., N)$ (Baron et al. 2005). If in Eq. (2) x is k-sparse signal and $\Psi = I$ then x can be found by solving (2) with $\ell 1$-minimization. If x itself is not sparse but has k-sparse representation, then (2) is transformed to:

$$y = \Phi\ \Psi\ c = \Theta\ c, \tag{3}$$

where Θ is so-called reconstruction matrix with dimensions $M*N$. Actually Eq. (3) is the same as (2) in cases when x is k-sparse, because then $\Psi = I$. Columns of Θ are called the holographic basis (Baron et al. 2005). Sparse representation c of signal x is found by means of holographic basis using $\ell 1$-minimization, but the signal x is found by means of (1). In the CS construction of well formed Φ is most important, there are restrictions to its content – the measurements in y should be independent and incoherent, so Φ must fulfill the restricted isometric property (RIP) and incoherence restrictions. The last restriction means, that the basis $\{\phi_m\}$ cannot sparsely represent the elements of the basis $\{\psi_n\}$ (incoherence of the two bases) (Donoho 2006; Candès et al. 2006a; Baron et al. 2005; Candès &

Tao 2006). RIP (initially called the *uniform uncertainty principle* in Candès et al. (2006a), Fornasier & Rauhut (2011) and Candès et al. (2006b)) is determined by orthogonality of Φ columns – less orthogonality demands more measurements for a complete recovery of x. If $\{\psi_n\}$ is some orthonormal basis, then the measurement matrix Φ obeys RIP if it is formed as

- N column vectors sampled uniformly at random on the unit sphere of $\mathbb{R}^M$;
- independent and identically distributed (i.i.d.) entries from the normal distribution with the mean value 0 and the variance $1/M$;
- i.i.d. entries from a symmetrical Bernoulli distribution $P(\Phi_{ij} = \pm\, 1/\sqrt{M}) = 1/2$;
- other sub-Gaussian distributions;
- the matrix elements using a special techniques (Candès & Walkin 2008; Xu et al. 2010).

2.2. Analysis of space signal classes

A rough analysis of the RA signal classes[2] shows that there are two basic classes of RA signals.

I. The radio emission of the following sources covers wide spectral ranges, usually more than 1 GHz: (1) most thermal emission-based radio sources with roughly blackbody or flat spectrum; (2) quasars and radio galaxies with the power-law spectral distribution; (3) pulsars with very steep spectrum (spectral indices $\alpha = -1.8$ to -1.6); (4) the Sun; (5) radio emission from auroras or lightning of planetary objects.

II. The emission of the following sources usually covers narrow spectral ranges: (1) spectral lines, (2) the emission from the heliopause triggered by an interplanetary shock, observable at kHz frequencies (Gurnett et al. 1993).

The wide-band signals are observed in channels (bands) with the bandwidths ranging from 1 to 32 MHz (Bezrukovs et al. 2012). After splitting in channels the spectrum of a wide band signal looks like white noise. From information theory it is known, that such signals are unpredictable, so incompressible. In the case when the level of the observed signal is below the system noise level, the signal on the input of a digitizing unit will be the white noise. Actually, because of filtering, the signal in the channels after registration will never be a real white noise, so it could be compressible but could not have a sparse representation. As the VIRAC participates in the ionospheric research and observations of space debris and asteroids, there are two more signal classes for analysis: the GPS signals used for the ionospheric research and the monochromatic 5 GHz signals emitted at Evpatoria (Crimea, Ukraine) by a planetary radar RT-70 ($P = 20$–100 kW) and reflected by near-Earth objects.

The bandwidth of the GPS signal in comparison with bandwidths of most of RA signals is quite narrow (Figure 4) – only about 20 MHz. Even more – for the ionospheric research only its central maximum is used, so the signal can be filtered to remove the noise and the remaining parts of the GPS signal spectrum. The GPS signal is powerful, so after filtering it should be well compressible.

Monochromatic signal recordings should be compressible very well if the signal,

[2] Kraus et al. (1986); Wilson et al. (2009); Burke & Graham-Smith (1996); Condon & Ransom (2007); Gurvits et al. (2005); Warwick (1967); Kellermann & Verschuur (1988).

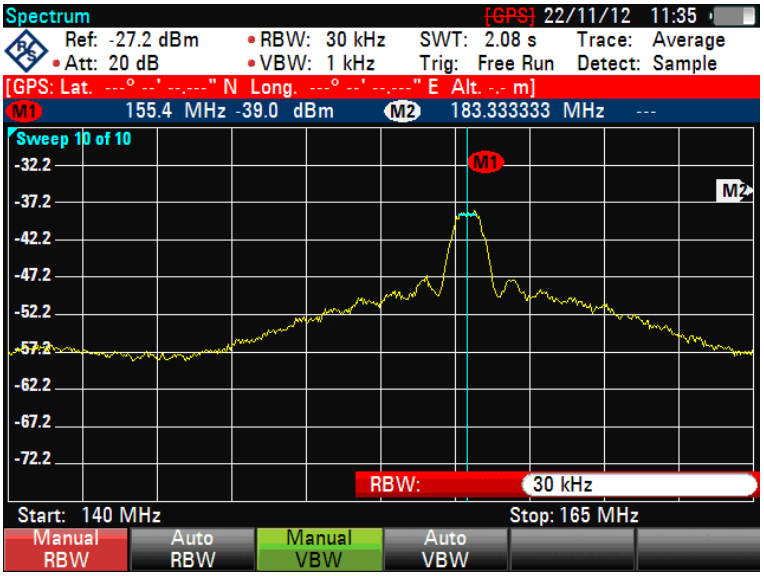


Fig. 4. The spectrum of a GPS signal.

reflected from the object, is sufficiently powerful (the noise makes the signal more unpredictable, so incompressible). As filtering decreases the signal informativity, it is expected that additional signal filtering will improve its compressibility.

2.3. Filtering

During this work two kinds of signal filtering have been done. In the text below one filtering is named pre-processing, another one – cleaning. Both filtering processes simply remove specific components of the signal spectrum as described below. The cleaning can be realized as described in Section 3.4.

2.3.1. Signal pre-processing

For the registration of the observed space signal the unit TN-16 (Bezrukovs et al. 2012) has been used. As can be seen from Figure 5a, this unit inserts into the captured signal so-called frequency marks, the signals with the frequencies $m \times 125$ kHz ($m = 1, 2,...$). For the analysis we took the signals with and without frequency marks – removing of these marks from the registered signal in the scope of this work is called 'signal pre-processing'. Pre-processing has been done by elimination of the spectrum components related to the frequency marks (see Figure 5b).

As it is seen from Figs. 5b, 6 and 7, the signal former can have a non-uniform frequency characteristic (the frequency-response curve) what could explain cyclic waves of spectral envelope of the pre-processed signals. A deeper analysis shows that frequency marks formed by TN-16 have side maxima which are not stable in frequency, so the filter of frequency marks has to be manually adjusted for each data file. In Tables 1–3, which give the results of analysis, for each class of signals two rows of the results are given – one row is for the unprocessed signals and the another row – for the pre-processed signals; rows are marked with the letters U and P, respectively.

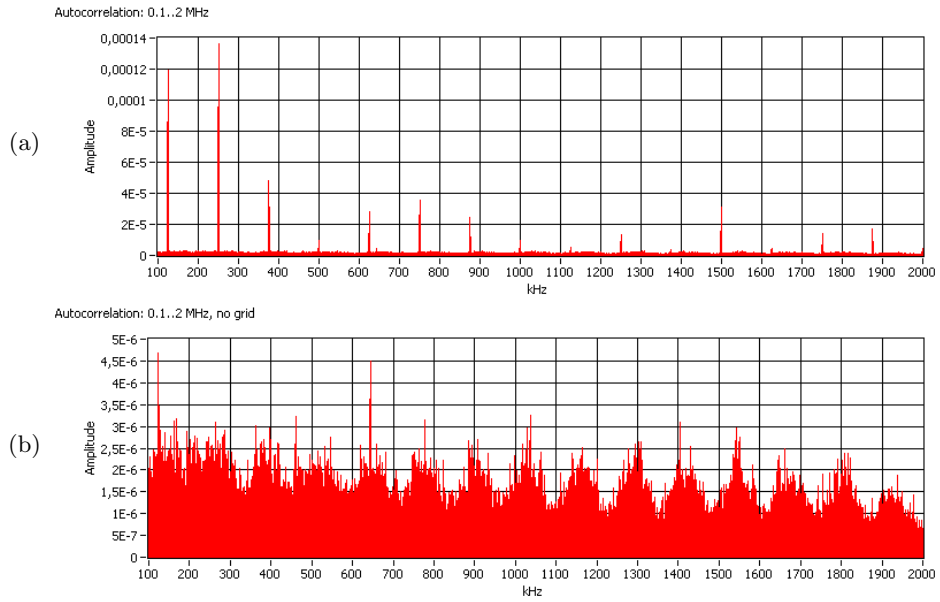


Fig. 5. The spectrum of the clear sky signal: (a) before pre-processing; (b) after pre-processing.

2.3.2. Signal cleaning

To check the effect of filtering, only the narrow-band signals with the signal level above the noise level can be taken for the analysis because: (1) for wide-band signals it is difficult to identify which components in the spectrum originate from space and which belong to the noise; (2) for narrow-band signals the filter frequency can be found only if the general spectrum component of the signal rises above the noise level (actually, in some cases the signal frequency can be roughly pre-calculated, in such cases it is possible to work with the signals of interest being below the noise level).

The pre-processed signal reflected by a space debris was filtered using a simple noise filter – DCT spectrum components with the magnitude below some level were zeroed (similar results will be given by the band-pass filter with the central frequency around the pre-calculated frequency of the reflected monochromatic signal). The noise filter thresholding level was chosen at 1% and 2% from the amplitude of the maximum amplitude of the spectrum component. The results, marked as F1 and F2, are given in the last two rows of Tables 1–3.

2.4. Space signal compressibility tests

The signal compressibility was tested using the three different approaches: (1) tests using special software made only for one purpose – compression of the files (file archivers); (2) entropy analysis of space signals; (3) analysis of space signals using the signal transforms.

2.4.1. Analysis of space signal compressibility using file archivers

Since our task was not testing all the archiving formats, only two formats have been checked: 'zip' and '7z'. Although the 'rar' archiving looks promising (Nieminen 2004), it was not used because of the proprietary compression algorithm [3]. To calculate the compression efficiency, the files containing space signals were compressed into the zip and 7z formats, then the sizes of the archived files were divided by the sizes of the uncompressed files to get the signal compression efficiency results comparable to the results given by the entropy analysis. Thus, in the scope of this work a lower value of the compression efficiency means better compression of the signal. If for some signal the compression efficiency is $\ll 1$, then the original signal should have the sparse representation. It is predictable that wide-band and noisy signals will have the compression efficiency close to 1, while the narrow-band and powerful monochromatic signals should have the compression efficiency $\ll 1$.

2.4.2. Entropy analysis of space signals

The entropy represents uncertainty or average informativity in bits of a message X (signal) per message symbol x (MacKay 2003):

$$H(X) = \sum_{x \in A_x} P(x) \log_2 P^{-1}(x), \tag{4}$$

where $P(x)$ is the probability to receive symbol x in message that can contain a set of symbols A_x; for $P(x) = 0$ there is defined $0 * \log_2(1/0) \equiv 0$. If $P(x)$ is uniform, the entropy is maximum (the signal is completely unpredictable):

$$H_{\max}(X) = \log_2 |A_x|, \tag{5}$$

where $|A_x|$ is the number of elements in the set of symbols A_x. The ratio

$$E(X) = H(X)/H_{\max}(X), \tag{6}$$

represents the efficiency of usage of the message space which shows, how much the original signal can be compressed – a lower $E(X)$ is for the redundant signals. To analyze the space signal redundancy, a line of virtual instruments (according LabVIEW terminology, VIs) in LabVIEW were created: VIs for calculation of the message entropy and the maximum entropy, VIs for analyzing the space data files (message symbol sizes 1, 2, 4, 8, 16 bits), and VI integrating all the previously mentioned.

2.4.3. Analysis of space signal compressibility using the signal transforms

To perform such test, a tool for space signal sparsity analysis in LabVIEW was created. It reads 1024 data samples from disk and finds a signal x transform (hopefully – sparse) c for the three cases: DCT, WT, HT. Then the transform coefficients are thresholded by a variable level and the recovered signal x is found with Eq. (1). Because the recovered signal is no more binary, it is thresholded against the average value of the recovered x. The PSNR of the final binary sig-

[3] http://en.wikipedia.org/wiki/Comparison-of-file-archivers

nal against the original signal is calculated. The amount of necessary spectrum components for perfect recovery directly represents the sparsity of the signal.

3. RESULTS AND DISCUSSION

To obtain space signals for the analysis, a dedicated observing session with the VIRAC-operated radio telescope RT-32 has been organized. Some data files were taken from the archive. The files of observational data were created by means of the unit TN-16 working with 4 MHz sampling frequency. The following datasets were analyzed: (1) the clear sky, (2) the supernova (SN) remnant 3C461/CasA, (3) the quasar 3C454.3, off the galaxy disk, (4) the CSS quasar 3C147.1, (5) Jupiter, (6) space debris, (7) the asteroid 90032 (2002 UM28) (radar mode, Molotov et al. 2008), (8) the quite Sun. Only the signals from (6) and (7) are narrow-band, others are wide-band.

3.1. Results of signal analysis using the file archivers

The results of the analysis are shown in Table 1. It is seen that the compression efficiency for the unprocessed data is close to 1 (93–99%). However, the processed data can be compressed also considerably – down to 60% of the original file size. The most compressed are the white noise signals – the quasar 3C 147.1 and the SN remnant 3C461/CasA signals (Figure 6). The spectra of both objects exhibit a pseudo white-noise with a cyclic spectrum characteristic curve. Between the assumed space signal classes, narrow-band signals (the space debris and asteroid observed in the radar mode) will be compressed better. However, Table 1 shows that these unprocessed or pre-processed signals do not compress well. This can be explained by a low signal power at the receiver input, when monochromatic radar signal is reflected back from a small object. As a result, the PSNR of the registered signal is low. The cleaned signal from space debris gives the compression efficiency 6.8% and 0.17% only. That means, that the signal has a sparse representation and the CS methods are applicable to them. In this case, the filtering increases sparsity of pre-processed signals up to about 450 times. Consequently, the filtering of the received signals is a key for success in application of the CS method in radio astronomy.

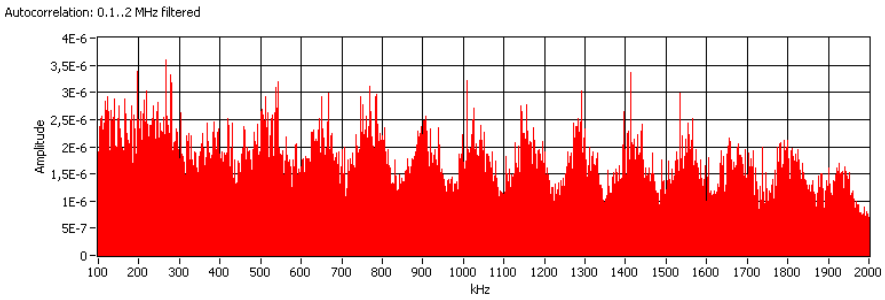
3.2. Results of the signal entropy analysis

For each dataset the entropy of available files was analyzed using a LabVIEW tool. Then the result was normalized to the maximum entropy for each symbol size to get the efficiency, Eq. (6). The results are given in Table 2, sorted by the results for SS=8. It is seen that the efficiency for the unprocessed data again is close to 1, but it is a little better than in the case of archivers test (92–99%). The best efficiency again is for the white noise signals – the CSS quasar 3C147.1 and Jupiter (Figs. 6 and 7). The efficiency increases with the size of the symbol, but for plausible analysis, for example, SS=8 and maximum entropy, the file should contain at least 256 symbols (bytes) – in such a case it is possible to find each symbol in the file at least once. For SS=16 the file should contain at least 2^{16} symbols or 2^{17} bytes = 128 kB. Usually symbols are not uniformly distributed, so for some statistics the file should contain more bytes. Because of the signal pre-processing and cleaning method limitations, all the analyzed files contain precisely

Table 1. Compression efficiency of RA signals.

Number	Object	Efficiency
4	Quasar 3C147.1 (P)	0.6043
2	SN remnant 3C461/CasA (P)	0.6141
1	Clear sky (P)	0.6192
3	Quasar 3C454.3 (P)	0.6196
5	Jupiter (P)	0.6222
6	Space debris (P)	0.7639
8	The Sun (P)	0.7879
7	Asteroid 90032 (P)	0.8537
7	Asteroid 90032 (U)	0.9275
8	The Sun (U)	0.9808
5	Jupiter (U)	0.9832
4	Quasar 3C147.1 (U)	0.9837
1	Clear sky (U)	0.9852
2	SN remnant 3C461/CasA (U)	0.9855
3	Quasar 3C454.3 (U)	0.9856
6	Space debris (U)	0.9865
6	Space debris (F1)	0.0680
6	Space debris (F2)	0.0017

(a)



(b)

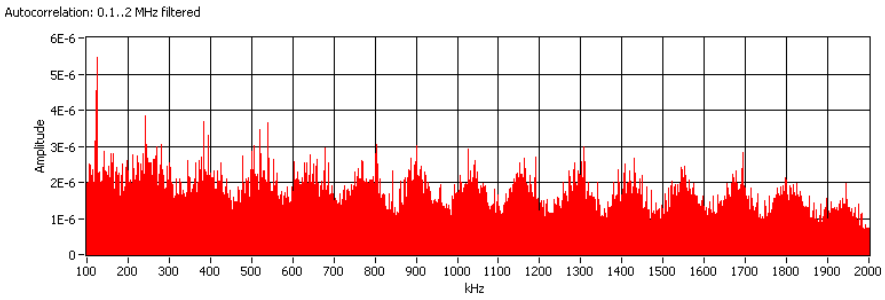


Fig. 6. The spectra of the two best compressible signals: (a) the CSS quasar 3C147.1; (b) the SN remnant 3C461/CasA.

2^{18} bytes of space signal data, so the efficiency calculated with SS=16 cannot be plausible – therefore in Table 2 we do not give the column for SS=16.

Table 2 shows that the best achievable efficiency at SS = 8 is 65%. This means that the modern archivers apply more effective algorithms for file compression,

Table 2. Message space usage efficiency of space signals for different message symbol sizes (SS).

No.	Object	SS=1	SS=2	SS=4	SS=8
4	Quasar 3C147.1 (P)	0.7025	0.6864	0.6698	0.6457
5	Jupiter (P)	0.6922	0.6813	0.6695	0.6528
2	SN remnant 3C461/CasA (P)	0.7112	0.6962	0.6813	0.6601
1	Clear sky (P)	0.7149	0.6997	0.6832	0.6607
3	Quasar 3C454.3 (P)	0.7177	0.7019	0.6857	0.6623
6	Space debris (P)	0.8104	0.8038	0.7963	0.7852
8	The Sun (P)	0.8039	0.7995	0.7956	0.7901
7	Asteroid 90032 (P)	0.8958	0.8920	0.8902	0.8874
7	Asteroid 90032 (U)	0.9411	0.9293	0.9228	0.9183
8	The Sun (U)	0.9911	0.9849	0.9798	0.9733
5	Jupiter (U)	0.9958	0.9873	0.9811	0.9751
4	Quasar 3C147.1 (U)	0.9963	0.9878	0.9816	0.9758
1	Clear sky (U)	0.9977	0.9893	0.9832	0.9774
2	SN remnant 3C461/CasA (U)	0.9976	0.9893	0.9832	0.9774
3	Quasar 3C454.3 (U)	0.9978	0.9895	0.9837	0.9779
6	Space debris (U)	0.9939	0.9891	0.9852	0.9794
6	Space debris (F1)	–	–	–	0.435
6	Space debris (F2)	–	–	–	0.395

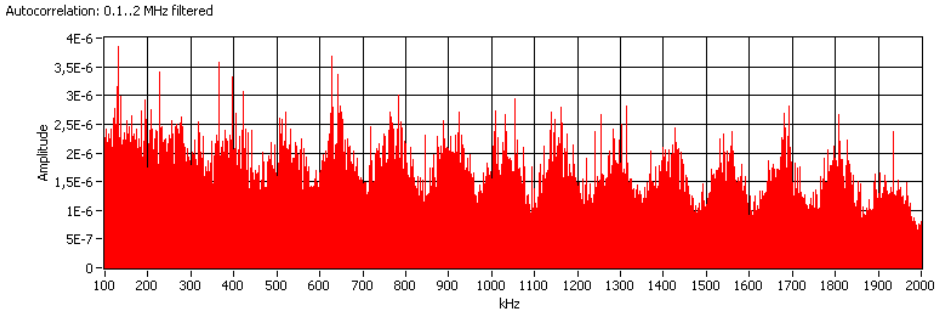


Fig. 7. The spectrum of the signal from Jupiter.

because in the previous subsection the best compression efficiency was 60%. The signals related to the noisy space debris and the asteroid in the radar mode do not compress well as expected. The entropy analysis of the space debris observations, cleaned from the noise, for SS = 8 gives the efficiency 43.5% and 39.5%. The comparison of the compression efficiency and the effectiveness of observations calculated by the simple entropy analysis shows that the results are quite similar to expected, and the simple entropy analysis is not a good tool for finding the signal compression efficiency – the new algorithms of archivers for compressible signals give better results.

3.3. Results using the signal transforms

The results of the signal transform analysis are sorted by the best result in Table 3. The results show, that even the unprocessed signals in the best case need only 569 ($\sim 56\%$) of the original spectrum components for a perfect signal recovery

Table 3. Results of space signal analysis using the signal transforms.

Obj#	Object	HT	WT	DCT	BEST
1	Clear sky (P)	767	477	314	314
5	Jupiter (P)	713	359	429	359
2	SN remnant 3C461/CasA (P)	741	471	434	434
3	Quasar 3C454.3 (P)	691	629	442	442
7	Asteroid 90032 (P)	777	533	569	533
8	The Sun (P)	752	674	564	564
1	Clear sky (U)	799	745	569	569
4	Quasar 3C147.1 (U)	821	696	580	580
6	Space debris (P)	755	582	590	582
4	Quasar 3C147.1 (P)	745	794	599	599
8	The Sun (U)	738	705	604	604
3	Quasar 3C454.3 (U)	868	691	630	630
6	Space debris (U)	830	676	636	636
7	Asteroid 90032 (U)	855	646	636	636
5	Jupiter (U)	805	647	661	647
2	SN remnant 3C461/CasA (U)	855	662	648	648
6	Space debris (F1)	–	312	–	312
6	Space debris (F2)	–	10	–	10

and 648 ($\sim$63%) in the worst case. With the pre-processed data in the best case only 314 ($\sim$31%) components from 1024 are needed for a perfect signal recovery (see the signal spectrum in Figure 5). This can be explained by thresholding done at the signal recovery stage. It is assumed that larger number of samples taken for the analysis gives better results – less spectral components (in %) for a perfect signal recovery will be needed for longer signals.

The sparsity analysis of the cleaned signal of the space debris gives that 312 spectrum components, and only 10 ($<1\%$) components are needed for a perfect recovery of signals if they are cleaned as described in Section 2.3.2. This again shows that the filtering (noise cleaning) of the observed signals is mandatory for CS applications for RA signals, and with CS achievable signal compression levels are comparable to those reached with the best file archivers.

3.4. Filtering in CS

As it was shown in previous chapters, filtering is a key for success of CS application on RA signals. The problem is that only signals from some source classes can be filtered to remove the system or background noise (cleaning):

(1) in the case of wide-band signal sources it is impossible to identify the noise and object signal components in the observed spectrum;

(2) when the signal is narrow-band and it is powerful enough to rise in the power spectrum above the noise level of the receiving system, it is possible to filter the noise and leave the narrow spectrum signal which can be handled using the CS approach, because the signal can become sparse;

(3) when the signal is narrow-band and its main frequency is known or can be predicted, it is possible to filter out most noise and leave the narrow spectrum signal and the remaining noise which can be handled using the CS approach since the filtering result can become sparse.

In the last two cases, CS has a hidden advantage described below. Usually, when the signal is filtered in the time domain using a filter kernel

$$k = \{k_1, k_2, ..., k_i\}, \tag{7}$$

it is possible to create a diagonal-like matrix

$$K = \begin{bmatrix} k_1 & k_2 & \cdots & k_i & 0 & 0 & \cdots & 0 \\ 0 & k_1 & \cdots & k_{i-1} & k_i & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & k_1 & \cdots & k_{i-1} & k_i \end{bmatrix} \tag{8}$$

with a kernel on the diagonal and get the filtered signal

$$f = Kx. \tag{9}$$

In the case of CS such signal sensing from Eq. (2) gives

$$y = \Phi\ f = \Phi\ Kx = Fx, \tag{10}$$

where $F = \Phi\ K$ is the filtering measurement matrix.

Thus, the filtering in CS can be done without dedicated filtering stage before taking measurements – F has to be pre-calculated for real time applications like CS of RA signals. Although the signal pre-processing is filtering too, the CS of such signal and the frequency mark removing following Eq. (10) should not be used for wide-band signals as after the frequency mark removing wide-band signals can be not sparse (see sections 3.1–3.3).

There can be the opposite use of Eq. (10) – in the cases when the measurement matrix Φ in Eq. (2) is poorly formed (Section 2.1.3; Candès et al. (2006a); Candès & Romberg (2007); Fornasier & Rauhut (2011)) and if it can be written (decomposed) as

$$\Phi = \Phi' K', \tag{11}$$

where Φ' is well formed and K' is a matrix with the structure like Eq. (8), then it is possible to find x sequentially solving Eq. (10) for f and Eq. (9) for x using the matrices Φ' and K'. Such way to find the signal x should give better results than a direct solving of Eq. (2) in the case when Φ is poorly formed because in this case the measurements actually are made using Eq. (10) – according to theory, the direct solving of Eq. (2) for x with badly formed Φ should give a worse result for the signal x than solving of Eq. (10) for f using the well formed matrix Φ'. This was not proved or tested here as it is not the aim of this paper.

4. FUTURE WORK

In the present study the following further research directions are identified.

(1) After filtering of the binary signals in spectral representation proper threshold level, for the conversion back to binary form should be used. In paper simple thresholding above the average level of the filtered signal is used to convert the signal back to the binary format. It has to be determined how the thresholding

scheme changes the filtering result;

(2) It has to be checked if larger data blocks and CS structures allow to reach better results in CS applications for RA;

(3) It has to be proved that the recovered signal can be improved following the signal recovery scheme described in Section 3.4.

5. CONCLUSIONS

(1) We showed that CS methods can be used to process binary RA signals by means of identity transform (Section 2.1.2).

(2) It was demonstrated that RA signals are compressible, usually they are neither sparse nor have sparse representations (Sections 3.1–3.3).

(3) It was demonstrated that RA signals with a noise-like spectrum can be well compressed – up to 60% from the initial file size (Sections 3.1–3.3).

(4) We showed that the filtering is a key for success of the application of CS methods to RA signals – after filtering RA signals can become sparse and the possible real-time CS signal compression level (up to 100 times, Section 3.3) can be comparable with the compression level that can be reached by the best file archivers with complicated and not real-time application (up to 500 times, Section 3.1).

(5) For narrow-band RA signals with good PSNR the CS method could be applicable even without filtering (compression level about 2 times, Section 3.3).

(6) For some wide-band RA signals the CS could be applicable too even without filtering (compression level up to 3 times – Section 3.3).

(7) We showed that no convolution in the filtering phase in case of CS is needed, filtering can be accomplished in the CS measurement phase thus decreasing the CS signal path complexity (Section 3.4).

(8) The analysis of the measurement matrices can help to improve the CS signal recovery.

(9) If in the measurement matrix a convolution filter is embedded, then the signal recovery can be done in two stage process that should give the restored signal with better quality than the signal after direct CS signal recovery (Section 3.4).

ACKNOWLEDGMENTS. The authors thank the European Regional Development Fund (ERDF), the Latvia-Lithuania Cross Border Cooperation Programme 2007–2013 project "JRTC Extension in Area of Development of Distributed Real-Time Signal Processing and Control Systems" (Cross-border DIS COS, LLIV-215) for the possibility to complete this scientific research. The publication of this article is supported by the European Regional Development Fund's project "International Competitiveness and Capacity-building of Satellite Research" (SATTEH, No.2010/0189/2DP/2.1.1.2.0/10/APIA/VIAA/019).

REFERENCES

Baron D., Duarte M. F., Sarvotham S. et al. 2005, in Proc. 45rd Conference on Communication, Control and Computing

Bezrukovs V., Shmeld I., Nechaeva M. et al. 2012, Latvian Journal of Physics and Technical Sciences, 6, 30

Burke B. F., Graham-Smith F. 1996, *An Introduction to Radio Astronomy*, Cambridge University Press

Candès E., Romberg J. 2007, Inverse Problems, 23, 969

Candès E. J., Romberg J., Tao T. 2006a, Information Theory, IEEE Transactions, 52, 489

Candès E. J., Romberg J. K., Tao T. 2006b Communications on Pure and Applied Mathematics, 59, 1207

Candès E. J., Tao T. 2006, Information Theory, IEEE Transactions, 52, 5406

Candès E. J., Wakin M. B. 2008, Signal Processing Magazine, IEEE, 25, 21

Condon J. J., Ransom S. M. 2007, *Essential Radio Astronomy Course*, http://www.cv.nrao.edu/course/astr534/ERA.shtml

Donoho D. L. 2006, Information Theory, IEEE Transactions, 52, 1289

Fornasier M., Rauhut H. 2011, *Compressive Sensing*, Handbook of Mathematical Methods in Imaging, Springer, p. 187

Gurnett D. A., Kurth W. S., Allendorf S. C., Poynter R. L. 1993, Science, 262, 199

Gurvits L. I., Frey S., Rawlings S. 2005, in *Radio Astronomy from Karl Jansky to Microjansky*, EAS Publication Series, vol. 15, EDP Sciences

Högbom J. A. 1974, A&AS, 15, 417

Kellermann K. I., Verschuur G. L. 1988, *Galactic and Extragalactic Radio Astronomy*, Springer

Kraus J. D., Tiuri M., Räisänen A. V., Carr T. D. 1986, *Radio Astronomy Receivers*, Cygnus-Quasar Books, Ohio

Li F., Cornwell T. J., de Hoog F. 2011, A&A, 528, A31

Lustig M., Donoho D., Pauly J. M. 2007, *Magnetic Resonance in Medicine*, Wiley Online Library, 58, 1182

MacKay D. J. C. 2003, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press

Nieminen J. 2004, *Archiver Comparison*, http://warp.povusers.org/

Suksmono A. B. 2009, in *Electrical Engineering and Informatics*, ICEEI, 1, 110

Warwick J. W. 1967, Space Science Reviews, 6, 841

Wenger S., Magnor M., Pihlström Y. et al. 2010, PASP, 122, 1367

Wiaux Y., Jacques L., Puy G. et al. 2009, MNRAS, 395, 1733

Wilson T. T. L., Rohlfs K., Hèuttemeister S. 2009, *Tools of Radio Astronomy*, Springer

Xu J., Pi Y., Cao Z. 2010, EURASIP Journal on Advances in Signal Processing, 2010, 43