Signal Processing 93 (2013) 2653–2663

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Performance analysis of partial segmented compressed sampling ☆

Wenbo Xu *, Yun Tian, Jiaru Lin

Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Article history:
Received 28 February 2012
Received in revised form
27 January 2013
Accepted 1 February 2013
Available online 10 February 2013

Keywords:
Analog-to-information conversion (AIC)
Compressed sampling
Segmented AIC
Restricted isometry property

ABSTRACT

Recently, a segmented AIC (S-AIC) structure that measures the analog signal by K parallel branches of mixers and integrators (BMIs) was proposed by Taheri and Vorobyov (2011). Each branch is characterized by a random sampling waveform and implements integration in several continuous and non-overlapping time segments. By permuting the subsamples collected by each segment at different BMIs, more than K samples can be generated. To reduce the complexity of the S-AIC, in this paper we propose a partial segmented AIC (PS-AIC) structure, where K branches are divided into J groups and each group, acting as an independent S-AIC, only works within a partial period that is non-overlapping in time. Our structure is inspired by the recent validation that block diagonal matrices satisfy the restricted isometry property (RIP). Using this fact, we prove that the equivalent measurement matrix of the PS-AIC satisfies the RIP when the number of samples exceeds a certain threshold. Furthermore, the recovery performance of the proposed scheme is developed, where the analytical results show its performance gain when compared with the conventional AIC. Simulations verify the effectiveness of the PS-AIC and the validity of our theoretical results.

1. Introduction

During recent years, a new theory of compressed sampling (CS) has emerged, which exploits the sparse prior to recover the signals from fewer samples than the number stated in the Nyquist theorem [1–4]. According to CS, these few samples are obtained by multiplying signals with a sampling matrix, which should satisfy the restricted isometry property (RIP) with overwhelming probability to guarantee reliable reconstruction [1,2].

To obtain compressed samples directly from an analog signal, analog-to-information conversion (AIC) has been proposed as a practical scheme [5–7]. The AIC structure consists of K parallel branches of mixers and integrators (BMIs), where each BMI multiplies the signal with a sampling waveform and the result is integrated over the sampling period T. In conventional AIC, the number of the collected samples equals to the number of BMIs. To get more samples, Taheri and Vorobyov [8] proposed the segmented AIC (S-AIC) scheme. The integration period T is divided into M equal segments such that K BMIs generate KM subsamples. By using the permuted results of these subsamples, this scheme can collect at most K^2 samples. The authors showed that the equivalent measurement matrix (EMM) of this scheme satisfies the RIP with overwhelming probability if the original matrix of BMI sampling waveforms satisfies it. Note that the EMM of the S-AIC is in fact a dense matrix, and the resulted AIC suffers from a higher hardware complexity when compared with the conventional AIC.

☆ Until the second round of review, a part of this work has been accepted as a short paper in International Journal of Electronics and Communications, 2012.

* Corresponding author. Tel./fax: +86 1062284023.
E-mail addresses: wenb.xu@gmail.com, xuwb@bupt.edu.cn (W. Xu).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.02.001

In this paper, we introduce a partial segmented AIC (PS-AIC) scheme, where each BMI only works within a partial time period and thus enjoys a reduced complexity. This scheme has been simply described in our recent short paper [9], where some preliminary results have been presented. PS-AIC is inspired by the favorable RIP feature of the block diagonal matrix (BDM). In this structure, K BMIs are divided into J groups that work in non-overlapping integration time ($J<K$), that is, the integration period in each BMI is reduced to T/J instead of T in S-AIC. Each group implements the same operations as those in S-AIC, and KM/J subsamples are collected from J groups. By using these subsamples in their original form and their permuted form, a larger number of samples (at most K^2/J samples) than the number of BMIs can be generated. We give the detailed proof for the RIP of the EMM of the PS-AIC,[1] which shows that the EMM satisfies the RIP when the number of samples is larger than a threshold. In addition, when the sparsity basis of the signal is the Fourier basis and the number of measurements satisfies a specific condition, EMM of the PS-AIC has the identical RIP condition as that of the S-AIC. Furthermore, we derive the mean squared error (MSE) of the PS-AIC scheme when empirical risk minimization method is used for recovery. The motivation for using this method is to make fair comparison with [8], which also develops MSE results when empirical risk minimization is adopted. Our analytical results show that PS-AIC enjoys better performance than the conventional AIC with K BMIs. Actually, PS-AIC can be implemented by using only K/J BMIs in the conventional AIC. Therefore, the proposed PS-AIC is a promising candidate for CS.

The rest of this paper is organized as follows. Section 2 presents the background of CS and S-AIC. The structure of the PS-AIC is described in Section 3, where we prove that the corresponding EMM satisfies the RIP. Section 4 provides the recovery performance analysis of PS-AIC. In Section 5, some simulation results are shown. Finally, Section 6 concludes this paper.

Throughout the paper, we denote vectors and matrices by boldface lowercase letters and boldface uppercase letters, respectively, e.g., $\mathbf{x}$ and $\mathbf{A}$. The lth element of $\mathbf{x}$ is denoted as x_l. The Euclidean norm of $\mathbf{x}$ is $\|\mathbf{x}\|_2=\sqrt{\mathbf{x}^H\mathbf{x}}$, $\|\mathbf{x}\|_1=\sum_l|x_l|$ is the l_1-norm, $\|\mathbf{x}\|_\infty=\max_l|x_l|$ is the l_∞-norm, and $\|\mathbf{x}\|_0$ denotes the number of nonzero entries in $\mathbf{x}$. For a matrix $\mathbf{A}$, $(\mathbf{A})_a^b$ denotes a submatrix containing the ath row to the bth row of $\mathbf{A}$, and $\mathbf{A}_{i,j}$ (or $(\mathbf{A})_{i,j}$) means the entry in $\mathbf{A}$. $\mathbf{A}^T$ stands for the transpose of $\mathbf{A}$. Furthermore, $\mathbf{A}=\mathrm{diag}(\hat{\mathbf{A}}_1,\hat{\mathbf{A}}_2,\ldots,\hat{\mathbf{A}}_J)$ is a block diagonal matrix with the jth block in the diagonal being $\hat{\mathbf{A}}_j$. We use $\Pr(\xi)$ to denote the probability of event ξ.

2. Background

2.1. Introduction of compressed sampling

Let $f(t)$ be an analog signal that can be represented as the linear combination of N basis functions $\{\psi_n(t)\}_{n=1}^N$ with $t\in[0,T]$. Specifically

$$f(t)=\sum_{n=1}^{N}x_n\psi_n(t)=\mathbf{x}^T\mathbf{\Psi}(t) \tag{1}$$

where $\mathbf{x}=(x_1,\ldots,x_N)^T$ is the coefficient vector and $\mathbf{\Psi}(t)=(\psi_1(t),\ldots,\psi_N(t))^T$. If $\mathbf{x}$ has S nonzero entries with $S\ll N$, $f(t)$ is called an S-sparse signal with $\{\psi_n(t)\}_{n=1}^N$ being the sparsity basis.

In the theory of compressed sampling, the sparse signal can be represented by its corresponding samples with a reduced dimension. One practical structure of deriving the samples is AIC shown in Fig. 1(a), which consists of K parallel BMIs. Each BMI is used to measure the analog signal against different random sampling waveforms $\phi_k(t)$ ($k=1,\ldots,K$), and the kth measurement can be obtained as

$$y_k=\int_0^T f(t)\phi_k(t)\,dt \tag{2}$$

Define the discrete equivalents of $\mathbf{\Psi}(t)$ and $\mathbf{\Phi}(t)=(\phi_1(t),\ldots,\phi_K(t))^T$ as $\mathbf{\Psi}\in\mathbb{R}^{N\times N}$ and $\mathbf{\Phi}\in\mathbb{R}^{K\times N}$ with the entries being $\psi_{m,n}=\int_{(n-1)T/N}^{nT/N}\psi_m(t)\,dt$ and $\phi_{k,n}=\int_{(n-1)T/N}^{nT/N}\phi_k(t)\,dt$, respectively [8], where $\mathbf{\Phi}$ is called the measurement matrix. Suppose that $\mathbf{\Psi}$ is an orthonormal matrix. Define the discrete counterpart of the signal $f(t)$ as $\mathbf{f}=\mathbf{\Psi}^T\mathbf{x}$. Using $\mathbf{\Phi}$, the noisy samples $\mathbf{y}=(y_1,\ldots,y_K)^T$ can be written as

$$\mathbf{y}=\mathbf{\Phi}\mathbf{f}+\mathbf{w} \tag{3}$$

where $\mathbf{w}$ is an identically and independently distributed (i.i.d.) Gaussian noise vector with entries of zero mean and variance σ^2.

In compressed sampling, a general condition for exact reconstruction is the RIP of the measurement matrix [8,10]. This is described in the following.

Definition 1. Assume that the matrix $\mathbf{\Phi}\in\mathbb{R}^{K\times N}$ consists of i.i.d. entries with variance $1/N$. We call that it satisfies the RIP of order S and conditioning $0<\delta<1$ if for every S-sparse signal $\mathbf{c}\in\mathbb{R}^N$, it has

$$\frac{K}{N}(1-\delta_S)\|\mathbf{c}\|_2^2\le\|\mathbf{\Phi}\mathbf{c}\|_2^2\le\frac{K}{N}(1+\delta_S)\|\mathbf{c}\|_2^2 \tag{4}$$

The recovery of the sparse signal from its samples $\mathbf{y}$ can be achieved by using different types of methods, see [11] for details. One method consists in considering convex optimization problems, e.g., the widely studied l_1-norm minimization [1]

$$\min\|\mathbf{x}\|_1 \quad \text{s.t. } \|\mathbf{\Phi}\mathbf{\Psi}\mathbf{x}-\mathbf{y}\|_2\le\eta \tag{5}$$

where η is the energy bound on $\mathbf{w}$ such that $\|\mathbf{w}\|_2\le\eta$. It is known that this method works well if $\mathbf{\Phi}$ satisfies the RIP.

In this paper, to compare the performance of the proposed scheme and S-AIC in [8], we consider the recovery method adopted in [8], i.e., the empirical risk minimization [12]. Define the risk of a candidate reconstruction $\hat{\mathbf{f}}$ to be $r(\hat{\mathbf{f}})=\|\hat{\mathbf{f}}-\mathbf{f}\|^2/N+\sigma^2$ and the empirical risk to be $\hat{r}(\hat{\mathbf{f}})=(1/K)\sum_{j=1}^K(y_j-\boldsymbol{\phi}_j\hat{\mathbf{f}})^2$, where $\boldsymbol{\phi}_j$ is the jth row of $\mathbf{\Phi}$. Then based on K samples, the candidate

[1] Xu et al. [9] only give an abridged version of the proof.