Digital Signal Processing 54 (2016) 12–26

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Sparse signal reconstruction via concave continuous piecewise linear programming ☆

Kuangyu Liu [a,*], Zhiming Xu [a,b], Xiangming Xi [a], Shuning Wang [a]

[a] Department of Automation, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China
[b] College of Science, Air Force Engineering University, Xi'an 710051, China

ARTICLE INFO

Article history:
Available online 6 April 2016

Keywords:
Compressed sensing
Continuous piecewise linear programming
Global optimization
γ valid cut

ABSTRACT

Compressed sensing (CS) is a new paradigm for acquiring sparse and compressible signals which can be approximated using much less information than their nominal dimension would suggest. In order to recover a signal from its compressive measurements, the conventional CS theory seeks the sparsest signal that agrees with the measurements via a great many algorithms, which usually solve merely an approximation of the l_0 norm minimization. In this paper, CS has been considered from a new perspective. We equivalently transform the l_0 norm minimization into a concave continuous piecewise linear programming based on the prior knowledge of sparsity, and propose a novel global optimization algorithm for it based on a sophisticated detour strategy and the γ valid cut theory. Numerical experiments demonstrate that our algorithm improves the best known number of measurements in the literature, relaxes the restrictions of the sensing matrix to some extent, and performs robustly in the noisy scenarios.

1. Introduction

Over the years, sparse representation in the signal processing attracts increasing attention. Signals have sparse coefficients in some transform domain, such as the wavelet or Fourier, and exploiting this sparsity offers significant advances in sensing system. The wisdom behind this is the compressibility of signals: most of the information contained in a signal resides in a few large coefficients. Therefore, in traditional applications one can first measure the entire signal, then merely retains the very small number of large coefficients. This intuitively raises the question: is it possible to perform a compact measurement directly on a sparse signal (measure only a part of the signal)? The fundamental results in [1–3] answer the above question affirmatively by introducing the theory of compressed sensing (CS).

CS provides an alternative to Shannon/Nyquist sampling for the acquisition of sparse or compressible signal that can be well approximated by just k ($\ll n$) components from an n-dimensional basis. In this framework one does not measure the n-dimensional signal directly, but rather inner products with m ($\ll n$) measurement vectors and then recovers the signal via certain reconstruction algorithms. CS integrates the signal acquisition and compression steps into a single process. In addition, the ratio m/n is very small, which offers the potential to simplify the sensing system. Hence, the implications of these facts are far-reaching, with applications in single-pixel camera [4], data compression [5], medical imaging [6], analog-to-digital converters [7,8], sensor networks [9,10], and so on.

The essential issue in the CS theory is the signal reconstruction. Although the recovery of the signal from the extremely limited measurements appears to be a severely ill-posed inverse problem, the prior knowledge of sparsity gives us solid hope for accurate reconstruction. Actually, the signal recovery can be achieved by searching for the sparsest one that agrees with the observed measurements.

Mathematically speaking, under the sparsity and noise-free assumptions, one can recover a k-sparse signal $\bar{x} \in \mathbb{R}^n$, namely $\| \bar{x} \|_0 \le k$ (e.g., the coefficient sequence of the signal in an appropriate basis), by solving the nonconvex optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \| x \|_0 \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{1}$$

where $\| \cdot \|_0$ denotes the l_0 "norm" that counts the number of nonzero elements, and the sensing matrix $A \in \mathbb{R}^{m\times n}$ is usually

☆ This project is jointly supported by the National Natural Science Foundation of China (61473165, 61134012) and the National Basic Research Program of China (2012CB720505).

* Corresponding author.
E-mail addresses: liu-ky12@mails.tsinghua.edu.cn (K. Liu), xzm13@mails.tsinghua.edu.cn (Z. Xu), xxm10@mails.tsinghua.edu.cn (X. Xi), swang@mail.tsinghua.edu.cn (S. Wang).

http://dx.doi.org/10.1016/j.dsp.2016.03.010
1051-2004/

generated by randomly sampling the columns independently from a certain distribution (e.g., the Gaussian distribution).

Unfortunately, problem (1) is known to be NP-hard and is generally impossible to be solved, as it usually requires to perform a combinatorial enumeration of all the feasible sparse situations. However, fundamental results in [2] show that a computationally tractable optimization problem yields an equivalent solution, which can be found by solving the basis pursuit (BP) problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \| x \|_1 \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{2}$$

as long as A satisfies the restricted isometry property (RIP). Problem (2) can be viewed as the closest convexification of problem (1) and is much more approachable, which can be easily solved with linear programming techniques. Shortly afterwards, a burst of researches in sparse signal reconstruction have been motivated by BP and RIP. More and more practical and sophisticated algorithms have been proposed, which can be categorized into the following five rough groups.

Optimization based algorithms: These approaches solve the convex or nonconvex programming problems whose minimizer is known to approximate the target signal. Obviously, aforementioned BP is of this kind. Kim et al. proposed a specialized interior-point method for solving l_1-regularized least squares problem

$$\min_x \| Ax - b \|_2^2 + \mu \| x \|_1, \mu > 0, \tag{3}$$

which uses the preconditioned conjugate gradients algorithm to compute the search direction [11]. Other developed methods for problem (3) include the Bregman iterative algorithms [12–14], the gradient projection method [15], and the shrinkage and subspace optimization [16]. Candes et al. described a method called iterative reweighted l_1 minimization (IRL1) consisting of solving a sequence of weighted l_1 minimization problems

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \textstyle\sum_{i=1} \omega_i |x_i| \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{4}$$

where the weights $\omega_1, \omega_2, \cdots, \omega_n$ are updated during each iteration based on the previous solution [17]. And this method requires fewer measurements than l_1 minimization. Wang and Yin presented an iterative support detection (ISD) method, which runs as fast as the best BP algorithms but requires significantly fewer measurements via solving convex truncated BP [18].

It is shown by Chartrand [19] that a nonconvex variant of BP could produce exact reconstruction with fewer measurements. Specifically, problem (2) is replaced by the l_p minimization,

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \| x \|_p^p \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{5}$$

where $\| x \|_p = (\sum_i |x_i|^p)^{1/p}$, $0 < p < 1$, is the l_p-quasi-norm of x. Then one can adopt a simple computational approach such as gradient descent with projection to compute local minimizers of problem (5). This work is extended and refined in the subsequent years. Chartrand considered the use of iteratively reweighted least squares (IRLS) approach for the above nonconvex problem [20–23], and the experiment results suggested that $p = 1/2$ seems suitable. Recently, researchers extended problem (5) to the matrix space, called M_p-minimization [24]. Regarding the nonconvex variants of BP, besides L_p-norm, log-sum function is also an effective sparsity-encouraging function which behaves very close to L_0-norm. Iterative reweighted methods and theoretical analysis based on log-sum minimization were studied in a number of works [25–27]. It should be noted that the nonconvexity of problem (5) means that all of the algorithms considered here are only designed to produce local minima. However, these local algorithms may give global solutions, if initialized by a point sufficiently close to the global optimum [19].

Iterative greedy algorithms: These methods build up an approximation of the signal by making locally optimal choices at each step. Gilbert et al. showed the way to incorporate greedy iterative strategies into fast sparse approximation algorithms and establish the first rigorous guarantees for greedy methods [28]. Tropp and Gilbert proved theoretically and empirically that orthogonal matching pursuit (OMP) is effective for CS [29]. Soon after, faster algorithms have been proposed, such as stagewise OMP (StOMP) [30], regularized OMP (ROMP) [31], compressive sampling matching pursuit (CoSaMP) [32], subspace pursuit (SP) [33], iterative hard thresholding (IHT) [34], accelerated iterative hard thresholding (AIHT) [35] and so on. The major advantages of this kind of algorithms are their fast speed and their ease of implementation.

Combinatorial algorithms: These methods acquire highly structured samples of the signal that support rapid reconstruction via group testing. Cormode and Muthukrishnan presented an approach of two sets of group tests with different separation properties that yields the first known polynomial time explicit construction of a non-adaptive transformation matrix and a reconstruction algorithm [36]. Gilbert et al. exhibited the chaining pursuit (CP) method which combines sublinear reconstruction time with stable and robust linear dimension reduction of all compressible signals [37]. However, simulations reveal that CP works well only when the signal is extremely sparse. Subsequently, Gilbert et al. presented heavy hitters on steroids (HHS) pursuit [38]. Unlike CP, HHS uses separate matrices for estimation, sifting, and noise reduction.

Statistics based algorithms: These methods connect CS to the statistical inference, which offers the potential for more precise estimation of signal or a reduction in the number of measurements. Ji et al. considered from a Bayesian perspective and utilized the relevance vector machine (RVM) for signal estimation [39]. Seeger and Nickisch extended these ideas to Bayesian experimental design and provided a approximate method based on expectation propagation [40]. Sarvotham et al. described a specific measurement scheme using an low density parity check like (LDPC-like) measurement matrix [41] or a CS-LDPC measurement matrix [42], and employed belief propagation techniques to accelerate the reconstruction of approximately sparse signals. The other related methods on application of Bayesian framework to sparse inverse problem can be found in [43] and the references therein.

Structured sparsity algorithms: These methods focus on a special kind of signals called structured sparsity models, which restrict the sparsity patterns of the approximations. Stojnic et al. developed an efficient recovery algorithm of block-sparse signals by minimizing a mixed l_2/l_1 norm which can be cast as a convex second-order cone programming [44] (see also Eldar and Mishali [45]). Baraniuk et al. introduced a model-based CS theory and proposed the model-based greedy algorithms for the recovery of the block-sparse signals and the tree structure signals [46–48]. Indyk and Price initiated a study of the recovery for the tree structure sparse signals under the Earth–Mover Distance [49]. By reducing the degrees of freedom of a signal, the structured sparsity models provide an immediate benefit to CS, which is a reduction in the number of measurements.

The major selling point for CS is that it uses a limited number of measurements to recover an entire sparse signal. We list some measurement requirements (relate to recovering the signal with high probability) of the aforementioned classic algorithms in Table 1, from which we can see that, the model-based approach uses the least number of measurements by requiring the maximal prior knowledge. In practical applications, one can easily obtain the prior knowledge about the numbers of nonzeros, but not the locations of nonzeros. Then a question naturally emerges: whether a