

ADAPTIVE COMPRESSED SENSING FOR VIDEO ACQUISITION

Hassan Mansour and Özgür Yılmaz

University of British Columbia, Vancouver - BC, Canada

ABSTRACT

In this paper, we propose an adaptive compressed sensing scheme that utilizes a support estimate to focus the measurements on the large valued coefficients of a *compressible* signal. We embed a “sparse-filtering” stage into the measurement matrix by weighting down the contribution of signal coefficients that are outside the support estimate. We present an application which can benefit from the proposed sampling scheme, namely, video compressive acquisition. We demonstrate that our proposed adaptive CS scheme results in a significant improvement in reconstruction quality compared with standard CS as well as adaptive recovery using weighted ℓ_1 minimization.

Index Terms— Compressed sensing, adaptive measurements, weighted acquisition, video acquisition

1. INTRODUCTION

Compressed sensing (CS) is a highly effective sub-Nyquist sampling paradigm for the acquisition of signals that admit sparse or nearly sparse representations in some transform domain. When the signal representation is strictly sparse, exact signal recovery can be achieved with sufficiently many measurements. However, when the representation is compressible (or nearly sparse), then the reconstruction error is proportional to the best k -term approximation of the signal, where k is the approximated sparsity level (see for e.g.: [1, 2, 3, 4]).

Let x be an arbitrary signal in \mathbb{R}^N and let x_k be its best k -term approximation with support T_0 . Suppose that we acquire $n \ll N$ measurements $y = Ax + e$, where A is an $n \times N$ measurement matrix and e is measurement noise with a two norm bounded by $\|e\|_2 \leq \epsilon$. Candés, Romberg and Tao [2] and Donoho [1] show that if $n \gtrsim k \log(N/k)$, then solving the following ℓ_1 minimization problem

$$\underset{u \in \mathbb{R}^N}{\text{minimize}} \quad \|u\|_1 \quad \text{subject to} \quad \|Au - y\|_2 \leq \epsilon \quad (1)$$

can stably and robustly recover x from y . In fact, it was shown in [2] that if x^* is the solution to (1) and A has the restricted

isometry property (RIP) with constant $\delta_{(a+1)k} < \frac{a-1}{a+1}$ for some $a > 1$, then

$$\|x^* - x\|_2 \leq C_0 \epsilon + \frac{C_1}{\sqrt{k}} \|x_{T_0^c}\|_1, \quad (2)$$

where C_0 and C_1 are well-behaved constants that depend on $\delta_{(a+1)k}$.

The above result shows that if $y = Ax$ and A has RIP with say $\delta_{3k} < \frac{1}{3}$ and x were a k -sparse signal, then recovery using ℓ_1 minimization would be exact. However, if x were compressible, then the recovery error (although stable) is proportional to $\|x_{T_0^c}\|_1$. On the other hand, given the same matrix A and if it is possible to acquire measurements $\tilde{y} = Ax_k$, then solving (1) using measurements \tilde{y} instead of y would recover a vector x^* such that

$$\|x^* - x_k\|_2 = 0 \quad \text{and} \quad \|x^* - x\|_2 = \|x_{T_0^c}\|_2. \quad (3)$$

Notice from (2) and (3) that under the same RIP conditions, solving (1) with measurements \tilde{y} results in an error bounded by the ℓ_2 norm of $x_{T_0^c}$ versus its ℓ_1 norm when y is used. The ℓ_1 norm of a long non-sparse vector is much larger than its ℓ_2 norm. In fact, their ratio could scale proportional to $\sqrt{N - k}$.

This observation raises the question of whether it is possible to sparsify a signal before compressively sampling it, a procedure which has an analogue in traditional sampling theory. In traditional Shannon-Nyquist sampling, a bandlimited analog signal can be reconstructed exactly from sufficiently many discretized measurements. However, in most practical applications, signals are not truly bandlimited. To prevent aliasing in the reconstruction, analog signals are first low pass filtered to limit the highest frequency content before acquiring the discrete measurements. Compressed sensing, on the other hand, deals with discrete signals that have sparse nonzero coefficients. In the case of non-sparse compressible signals, we propose to use “sparse-filtering” (as opposed to low pass filtering) before compressive sampling. However, pre-filtering would require the acquisition of the full signal to find the locations of the largest coefficients, information that is generally not available *a priori*. Fortunately, in many applications such as audio and video, it is possible to draw an estimate of the support of the largest coefficients before acquiring the signal.

In this paper, we propose an adaptive compressed sensing scheme that utilizes a support estimate to focus the measurements on the large valued coefficients of a *compressible*

hassanm@cs.ubc.ca; oyilmaz@math.ubc.ca

Both authors were supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant DNOISE II (375142-08). Ö. Yılmaz was also supported in part by an NSERC Discovery Grant.

signal. Our scheme is related to the works by [5] and [6] and motivated by the previous work presented in section 2. Our scheme, presented in section 3, embeds a sparse-filtering stage into the measurement matrix by weighting down the contribution of signal coefficients that are outside the support estimate with small nonzero weights $\omega < 1$. We then illustrate how adaptive CS can be applied to compressive video acquisition. Finally, we demonstrate through numerical simulations in section 4 that our proposed adaptive CS scheme results in a significant improvement in reconstruction quality compared with standard CS as well as adaptive recovery using weighted ℓ_1 minimization.

2. RELATED WORK

Our proposed scheme is related to the works in [5] and [6]. In [5], a data-adaptive procedure is proposed that utilizes information from previous observations to focus subsequent measurements into subspaces that are increasingly likely to contain true signal components. In [6], a variable density sampling scheme is proposed in which a sampling profile is chosen to minimize the mutual coherence between the sampling operator and the sparsity basis. Our scheme differs from the above works in that first, it deals with compressible signals and second, it focuses the measurements onto the high valued transform coefficients of the signal by “sparse-filtering” the signal before acquisition.

Our work is motivated by recent results on adaptive recovery from standard CS measurements using weighted ℓ_1 minimization [7]. Given a support estimate set \tilde{T} , the weighted ℓ_1 minimization problem is defined as

$$\underset{u}{\text{minimize}} \quad \|u\|_{1,w} \text{ subject to } \|Au - y\|_2 \leq \epsilon \quad (4)$$

where the weighted ℓ_1 norm $\|u\|_{1,w} = \sum_{i=1}^N w_i |x_i|$, and the weight vector w is given by

$$w_i = \begin{cases} \omega, & i \in \tilde{T}, \\ 1, & i \in \tilde{T}^c. \end{cases}$$

Denote by $\alpha = \frac{|\tilde{T} \cap T_0|}{|\tilde{T}|}$ the accuracy of \tilde{T} with respect to T_0 . It was shown in [7] that if $\alpha > 0.5$ and the matrix A satisfies RIP with constant $\delta_{(a+1)k} < \frac{a-\gamma}{a+\gamma}$ for $\gamma = \gamma(\alpha, \omega) < 1$ for some $a > 1$, then

$$\|x_w^* - x\|_2 \leq C_0(\alpha)\epsilon + C_1(\alpha)k^{-1/2} \left(\omega \|x_{\tilde{T} \cap T_0^c}\|_1 + \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right), \quad (5)$$

where $C_0(\alpha)$ and $C_1(\alpha)$ are well-behaved constants that depend on the measurement matrix A , the weight ω , and the parameters α and ρ . In fact, if $\alpha > 0.5$, then the error bound constants $C_0(\alpha)$ and $C_1(\alpha)$ are smaller than their corresponding constants in the case of standard ℓ_1 minimization (1). Moreover, empirical studies in [7] have shown that a value

$\omega \approx 0.5$ results in the most reliable recovery especially when the accuracy of the support estimate is not known.

The recovery from compressed sensing measurements does not only depend on the sparsity of the signal but also on the relative decay of the transform coefficients. The following proposition from [8] relates the support recovery capabilities of ℓ_1 minimization and the decay of a compressible signal.

Proposition 2.1 ([8]). *Suppose that A has the null space property (NSP) [4] of order k with constant c_0 and*

$$\min_{j \in S} |x(j)| \geq (\eta + 1) \|x_{T_0^c}\|_1, \quad (6)$$

where $\eta = \frac{2c_0}{2-c_0}$. Then for some $s \leq k$, the set $S \subseteq \tilde{T}$, where $S = \text{supp}(x_s)$ and \tilde{T} is the support of the largest k -coefficients of the solution to (1).

The proposition states that if the tail of the sorted coefficients of the signal x decays fast enough, then ℓ_1 minimization is guaranteed to recover the support of at least the largest s coefficients of x . Such a support estimate can then be used when sampling other signals with a similar support. Moreover, the proposition suggests that if we are able to attenuate the tail of these similar support signals, then ℓ_1 minimization should recover a larger portion of the support of their large coefficients.

3. ADAPTIVE COMPRESSED SENSING OF COMPRESSIBLE SIGNALS

In this section we formulate the adaptive compressed sensing scheme and describe its application to video compressed sensing.

3.1. Scheme description

As discussed in the introduction, we want to recovery a signal $x \in \mathbb{R}^N$ from $n \ll N$ compressive and noisy measurements $y = Ax + e$. We assume that x is a compressible signal and D is an orthonormal basis that spans the space of x . Suppose that we have a (possibly inaccurate) estimate \tilde{T} of the support of the largest ρk coefficients of x for some $\rho \in \mathbb{R}_+$, and denote by T_0 the support of the best k -term approximation of x .

Contrary to the weighted ℓ_1 approach which mixes the high and low valued coefficients equally into the measurement vector, we wish to downplay the contribution of the low valued coefficients. For that purpose, we define a weighting vector $\tilde{w} \in \mathbb{R}^N$ such that

$$\tilde{w}_i = \begin{cases} 1, & i \in \tilde{T} \\ \omega, & i \in \tilde{T}^c, \end{cases} \quad (7)$$

where $0 < \omega < 1$, and the weighting matrix $W \in \mathbb{R}^{N \times N}$ such that

$$W = \text{Diag}(\tilde{w}). \quad (8)$$

Define the target “filtered” representation z of the signal x as

$$z = Wx. \quad (9)$$

It is easy to see that z and x have the following relation

$$\begin{cases} z_{\tilde{T}} = x_{\tilde{T}}, \\ z_{\tilde{T}^c} = \omega x_{\tilde{T}^c}. \end{cases} \quad (10)$$

We are now interested in recovering the vector z from the adapted measurements $\tilde{y} = AWx + e$. Notice that for small ω and reasonably accurate \tilde{T} , the sorted coefficients of z will have a smaller tail than those of x . We could solve the ℓ_1 minimization problem with measurements \tilde{y} :

$$\underset{u \in \mathbb{R}^N}{\text{minimize}} \quad \|u\|_1 \text{ subject to } \|Au - \tilde{y}\|_2 \leq \epsilon$$

to recover an approximation of $u^\#$. However, the weighted ℓ_1 results in [7] indicate that since a support estimate is known, solving the weighted ℓ_1 problem results in a better approximation of the signal than standard ℓ_1 minimization.

Consequently, we define the adaptive compressed sensing (adaptive CS) problem as follows

$$\underset{z \in \mathbb{R}^N}{\text{minimize}} \quad \|z\|_{1, \tilde{w}^{-1}} \text{ subject to } \|Az - \tilde{y}\|_2 \leq \epsilon, \quad (11)$$

which results in the target filtered vector $z^\#$, the sparse approximation of z .

Remark 3.1. In order to remain in the compressed sensing setting, the set \tilde{T} would have to be of size $|\tilde{T}| > n$. Otherwise, one should simply acquire $x_{\tilde{T}}$, i.e. set $\omega = 0$. However, this approach fails when the support of the signal we are measuring evolves over time, as in the case of video signals. In such situations, using $|\tilde{T}| \leq n$ and setting $\omega > 0$ allows $|\tilde{T}_j|$ to vary according to the new measurements of frame j . This prevents from getting stuck with the support of the first frame.

3.2. Application to video acquisition

One natural application for the adaptive CS scheme is video compressed sensing. Traditional video acquisition techniques capture a full frame (or image) in the pixel domain at a specific frame rate. The number of pixels acquired per image defines the spatial sampling rate, while the number of frames acquired per second defines the temporal sampling rate. Since the temporal sampling rate is usually high, a group of adjacent video frames are temporally correlated which is reflected in their spatial transform coefficients having large valued entries in roughly the same locations.

Suppose that the video signal is partitioned into m data blocks or frames f_j , $j \in \{1 \dots m\}$. For the first block f_1 , no prior information is available and an approximation $f_1^\#$ of the block is recovered from random measurements $y_1 = Rf_1$ using standard ℓ_1 minimization, where R is a random restriction matrix. For the subsequent data blocks f_j , $j \in$

$\{2 \dots m\}$, a support estimate \tilde{T}_{j-1} is first extracted from the transform coefficients of the recovered $(j-1)$ th data block. We then build the weighting matrix W_j according to (8) and generate a “sparsifying” matrix $M_j = D^H W_j D$. The new measurements \tilde{y}_j are adapted such that $\tilde{y}_j = RM_j f_j$ and the corresponding data block is recovered using adaptive CS (11) with $A = RD^H$ resulting in the approximation $f_j^\# = D^H z^\#$.

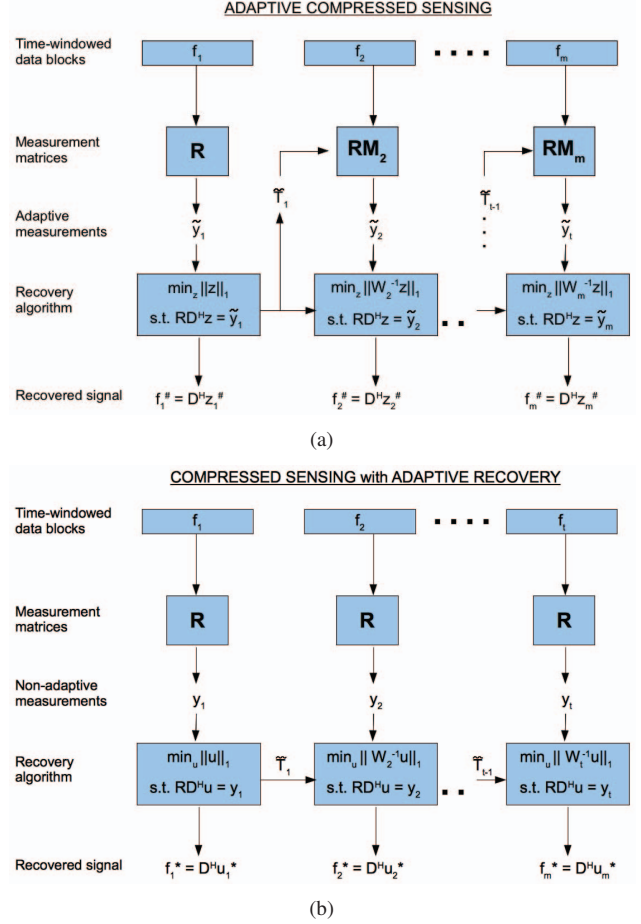


Fig. 1. Block diagrams illustrating the difference in acquisition and recovery between (a) the adaptive compressed sensing scheme and (b) the standard compressed sensing scheme with adaptive recovery using weighted ℓ_1 minimization. For every data block $j \in \{2 \dots t\}$, a support set \tilde{T}_j is identified from the previous block and a weighting matrix W_j is generated with weights equal to 1 and ω applied to the sets \tilde{T}_j and \tilde{T}_j^c , respectively. The sparsifying filters are $M_j = D^H W_j D$.

Figure 1 illustrates the difference in sampling structure and the recovery algorithm between the proposed adaptive CS and standard CS with adaptive recovery using weighted ℓ_1 minimization. The data blocks represent video frames. It can be seen that the main difference between the two sampling schemes lies in utilizing the support estimate in the measurement process in the case of adaptive CS.

Table 1. List of the sampling schemes and recovery algorithms

Sampling scheme	Support estimate size	Measurement matrix	Recovery algorithm
Standard CS	—	R	$\min_u \ u\ _1 \quad \text{s.t.} \quad RD^H u = Rf$
Weighted recovery	$\frac{n}{2} / \log(N/n)$	R	$\min_u \ u\ _{1,w^{-1}} \quad \text{s.t.} \quad RD^H u = Rf$
Adaptive CS	$0.7n$	$RD^H WD$	$\min_u \ u\ _{1,w^{-1}} \quad \text{s.t.} \quad RD^H u = RD^H WDf$

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the effectiveness of our adaptive CS scheme in recovering compressible signals by applying it to video sequences. We compare the recovery performance with that of standard CS using ℓ_1 minimization and adaptive recovery using weighted ℓ_1 minimization.

Table 1 lists the sampling parameters and recovery algorithms associated with each scheme. Note that the support estimate size indicated in the table is the empirically optimal size for each scheme. We consider two reference video sequences: Foreman and Mobile¹, at QCIF (176 × 144) spatial resolution. From each video frame, $n = N/4$ measurements are acquired, where $N = 25344$ is the number of pixels in the recovered video frame.

In the case of standard CS and adaptive (weighted ℓ_1) recovery, the measurements are simply the readings of n randomly chosen pixels in the CCD/CMOS sensor array. Denote by R the sampling operator that chooses these n pixels from the N pixels. On the other hand, the adaptive CS measurements are acquired by multiplying the entire frame with the matrix RM , where M is the sparse-filtering matrix. In particular, we choose $M = D^H W^2 D$, where D is the two dimensional discrete cosine transform (DCT) matrix and W is a diagonal weighting matrix as defined in (8) with $\omega = 0.5$. Figure 2 shows the recovered SNRs per video frame of the Foreman and Mobile sequences using each of the schemes in Table 1. The figure shows that for the Foreman sequence, adaptive CS achieves an average of 3.7dB gain over standard ℓ_1 and 2.73dB gain over weighted ℓ_1 minimization. For the Mobile sequence, the average improvement is 2.3dB over standard ℓ_1 and 1.75dB over weighted ℓ_1 .

5. REFERENCES

- [1] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, 2006.
- [3] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, 2006.

¹The reference video sequences Foreman and Mobile are available from: <http://trace.eas.asu.edu/yuv/>.

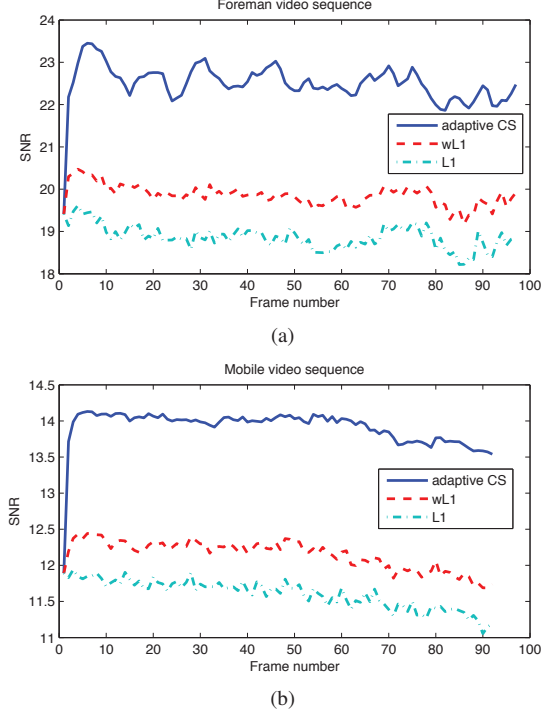


Fig. 2. Comparison of the recovered signal to noise ratio (SNR) for (a) the Foreman sequence and (b) the Mobile sequence between adaptive CS, standard CS, and standard CS with adaptive recovery (weighted ℓ_1 minimization).

- [4] A. Cohen, W. Dahmen, and R. DeVore, “Compressed sensing and best k-term approximation,” *Journal of the American Mathematical Society*, vol. 22, no. 1, pp. 211–231, 2009.
- [5] J. Haupt, R. Nowak, and R. Castro, “Adaptive sensing for sparse signal recovery,” in *IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009.*, jan. 2009, pp. 702–707.
- [6] G. Puy, P. Vandergheynst, and Y. Wiaux, “On variable density compressive sampling,” *Signal Processing Letters, IEEE*, vol. 18, no. 10, pp. 595–598, oct. 2011.
- [7] M. P. Friedlander, H. Mansour, R. Saab, and Ö. Yılmaz, “Recovering compressively sampled signals using partial support information,” *to appear in the IEEE Trans. on Inf. Theory*.
- [8] H. Mansour and Ö. Yılmaz, “Weighted ℓ_1 minimization with multiple weighting sets,” in *SPIE, Wavelets and Sparsity XIV*, August 2011.