COMPRESSIVE VIDEO RECOVERY WITH UPPER AND LOWER BOUND CONSTRAINTS

David R. Jones, Rachel O. Schlick, and Roummel F. Marcia

Department of Applied Mathematics, University of California, Merced, CA USA

ABSTRACT

The recovery of sparse images from noisy, blurry, and potentially low-dimensional observations can be accomplished by solving an optimization problem that minimizes the least-squares error in data fidelity with a sparsity-promoting regularization term (the so-called $\ell_2 - \ell_1$ minimization problem). This paper focuses on the reconstruction of a video sequence of images where known pixel-intensity bounds exist at each video frame. It has been established that the $\ell_2 - \ell_1$ minimization problem can be solved efficiently using gradient projection, which was recently extended to solve general boundconstrained $\ell_2 - \ell_1$ minimization problems. Furthermore, the video reconstruction can be made more efficient by exploiting similarities between consecutive frames. In this paper, we propose a method for reconstructing a video sequence that takes advantage of the interframe correlations while constraining the solution to satisfy known a priori bounds, offering a higher potential for increasingly accurate reconstructions. To demonstrate the effectiveness of this approach, we have included the results of our numerical experiments.

Index Terms— Video signal processing, optimization, gradient methods, compressed sensing

1. INTRODUCTION

The latest advances in video reconstruction take advantage of the assumption that the signal of interest (denoted f^*) is sparse or compressible within some basis or representation. While this has been widely known to hold for a variety of natural signals, studies in the field of compressed sensing [1, 2] have provided substantial theoretical support that employing sparsity constraints in signal reconstruction algorithms can lead to more accurate reconstructions [3]. In addition to sparsity, utilizing any additional information already known about f^* can assist in procuring increased accuracy in our estimates with minimal changes in the overall computational cost [4]. Recent advances have also indicated that using the correlation between frames of a video, assuming that the changes between frames is small, can also boost the accuracy of the reconstruction [5]. In this paper, we account for the boundary constraints on f, i.e., $b_U \ge f \ge b_L$, where b_U and b_L correspond to the upper and lower bounds on f respectively, while simultaneously incorporating inter-frame correlation to provide even greater accuracy in our reconstructions. This is accomplished through the development of a gradient projection method that can efficiently and accurately solve the constrained video reconstruction problem.

2. PROBLEM FORMULATION

We will consider the problem of approximating a series of images comprising a video $\{f_t^*\}_{t=1}^T \in \mathbb{R}^n$ by observing the linear projection $\{y_t\}_{t=1}^T \in \mathbb{R}^m$ where f_t is the *t*-th frame in a video that is made up of T sequential images. This can be formulated as $y_t = A_t f_t^* + \eta_t$ where $A_t \in \mathbb{R}^{mxn}$ projects the image linearly onto a *m*-dimensional set of observations and $\eta_t \in \mathbb{R}^m$ is an error vector associated with sensor noise, quantization errors, etc. In [5], we gave a detailed outline of applying gradient projection to solve twoframe and four-frame non-negative $\ell_2 - \ell_1$ minimization problem. In [4], we detailed how to impose additional constraints (as many applications have maximum and minimum values of $\{f_t^*\}$) to the single-frame $\ell_2 - \ell_1$ minimization problem. Here, we propose an efficient solution to this $\ell_2 - \ell_1$ sparse video minimization problem combining these two methods.

2.1. $\ell_2 - \ell_1$ Minimization

Through the use of an orthonormal basis W, we can create a sparse (or mostly zeroes) vector $\{\theta_t^*\}$, where $f_t^* = W\theta_t^*$. Common methods in current literature solve the following convex $\ell_2 - \ell_1$ optimization problem for each time frame t:

$$\widehat{\theta_t} \equiv \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \quad \frac{1}{2} \| y_t - A_t W \theta_t \|_2^2 + \tau \| \theta_t \|_1 \qquad (1)$$

subject to $b_U \ge W \theta_t \ge b_L$

where $\tau > 0$ is a regularization parameter and $b_U, b_L \in \mathbb{R}^n$ are the vectors of upper and lower bounds on $\{f_t^*\}$ respectively. The accuracy of our estimates can be improved by taking advantage of the relationships between sequential frames, as described below.

2.2. Video Reconstruction

In a video where each frame changes only slightly from the previous frame, the reconstruction of a frame is usually a good approximation for the subsequent frame (i.e. $f_t^* \approx f_{t+1}^*$ and therefore $\theta_t^* \approx \theta_{t+1}^*$). However, rather than solving for θ_t^* and θ_{t+1}^* , we solve for θ_t^* and $\Delta \theta_t^* = \theta_{t+1}^* - \theta_t^*$, which will be much sparser than θ_{t+1}^* . This approach results in the coupled optimization problem:

$$\begin{bmatrix} \widehat{\theta}_{t} \\ \Delta \widehat{\theta}_{t} \end{bmatrix} = \underset{\theta_{t}, \Delta \theta_{t} \in \mathbb{R}^{n}}{\operatorname{argmin}} \frac{1}{2} \left\| \begin{bmatrix} y_{t} \\ y_{t+1} \end{bmatrix} - A_{t,2} \overline{W}_{2} \begin{bmatrix} \theta_{t} \\ \Delta \theta_{t} \end{bmatrix} \right\|_{2}^{2} + \tau_{1} \|\theta_{t}\|_{1} + \tau_{2} \|\Delta \theta_{t}\|_{1}$$
subject to
$$b_{U} \geq W \theta_{t} \geq b_{L},$$

$$b_{U} \geq W (\theta_{t} + \Delta \theta_{t}) \geq b_{L} \qquad (2)$$

where $A_{t,2} = \begin{bmatrix} A_t & 0 \\ 0 & A_{t+1} \end{bmatrix}$, $\overline{W}_2 = \begin{bmatrix} W & 0 \\ W & W \end{bmatrix}$, and $\tau_1, \tau_2 > 0$. We can extend this approach to solve multiple frames simultaneously. In previous work on videos, we observed that when we held

This work was supported by NSF Grant DMS-09-65711 and by the University of California Leadership Excellence Through Advanced Degrees program.

constant the amount of processing time allotted per frame, the accuracy generally increases with each additional frame processed [6]. It must be noted, however, that the improvement in accuracy diminishes when the size of the problem is such that very few reconstruction iterations are run within the allotted time.

3. BOUND CONSTRAINED GRADIENT PROJECTION

Solving the unconstrained $\ell_2 - \ell_1$ minimization problem (1) (or its equivalent variants) can be done in various ways. Gradient Projection for Sparse Reconstruction (GPSR) is a method that performs particularly well [7]. Here, we describe how this gradient projection approach can be applied to solve a general upper and lower bounded multi-frame optimization problem. In [4], we outlined in detail how we apply gradient projection to solve the single-frame upper and lower bounded $\ell_2 - \ell_1$ minimization problem (1). In this section, we describe our proposed method for efficiently solving for two frames (2) simultaneously. We note that this approach can be extended to systems with arbitrarily many frames.

3.1. Gradient Projection

To use gradient-based methods to solve (2), we decompose θ_t , $\Delta \theta_t$ into there positive and negative components to reformulate the objective function so that it is differentiable: $\theta_t = u_t - v_t$, $\Delta \theta_t = \Delta u_t - \Delta v_t$, with $u_t, v_t, \Delta u_t, \Delta v_t \ge 0$ and $u_t^T v_t = \Delta u_t^T \Delta v_t = 0$ so that $\|\theta_t\|_1 = \mathbb{1}_n^T (u_t + v_t)$ and $\|\Delta \theta_t\|_1 = \mathbb{1}_n^T (\Delta u_t + \Delta v_t)$ where $\mathbb{1}_n \in \mathbb{R}^n$ is the *n*-vector of ones. We can now rewrite (2) as

where the solution is givn by $\hat{u}_t, \hat{v}_t, \Delta \hat{u}_t, \Delta \hat{v}_t$ and where

$$\Phi(u_t, v_t, \Delta u_t, \Delta v_t) = \frac{1}{2} \left\| \begin{bmatrix} y_t \\ y_{t+1} \end{bmatrix} - A_{t,2} \overline{W}_2 \begin{bmatrix} u_t - v_t \\ \Delta u_t - \Delta v_t \end{bmatrix} \right\|_2^2 + \tau_1 \mathbb{1}_n^T (u_t + v_t) + \tau_2 \mathbb{1}_n^T (\Delta u_t + \Delta v_t)$$
(4)

The reconstruction is now given by $\hat{f}_t = W(\hat{u}_t - \hat{v}_t)$, and $\hat{f}_{t+1}^{(0)} = W((\hat{u}_t - \hat{v}_t) + (\Delta \hat{u}_t - \Delta \hat{v}_t))$ is used to initialize the next optimization problem. We note that the penalty terms will naturally enforce the conditions $u_t^T v_t = \Delta u_t^T \Delta v_t = 0$.

We can simplify our notation by first letting $e_{j,k}$ be the row vector of length k with one in the first j entries and zero everywhere else. Using the Kronecker product \otimes , we define the following:

$$\widetilde{W_j} = e_{j,2} \otimes \begin{bmatrix} W & -W \end{bmatrix}$$
 and $\widetilde{I_j} = e_{j,2} \otimes \begin{bmatrix} I & -I \end{bmatrix}$.

We then let $\widetilde{W} = [\widetilde{W}_1; \widetilde{W}_2]$. and $z_t = [u_t; v_t; \Delta u_t; \Delta v_t]$, where $\widetilde{W} \in \mathbb{R}^{2nx4n}$ and $z_t \in \mathbb{R}^{4n}$. We can now write (3) compactly as

$$\widehat{z}_{t} \equiv \underset{z \in \mathbb{R}^{4n}}{\operatorname{argmin}} \quad \frac{1}{2} \left\| \begin{bmatrix} y_{t} \\ y_{t+1} \end{bmatrix} - A_{t,2} \widetilde{W} z_{t} \right\|_{2}^{2} + \begin{bmatrix} \tau_{1} \mathbb{1}_{2n} \\ \tau_{2} \mathbb{1}_{2n} \end{bmatrix}^{T} z_{t}$$

subject to $Cz_{t} \geq d$

where the resconstruction is given by $\left[\widehat{f}_t; \widehat{f}_{t+1}\right] = \widetilde{W} z_t$ and where

$$C = \left[I_{4n}; \widetilde{W}_1; -\widetilde{W}_1; \widetilde{W}_2; -\widetilde{W}_2 \right] \in \mathbb{R}^{8nx4n}$$

$$d = \left[0_{4n}; b_2; b_2 \right] \in \mathbb{R}^{8n}, \quad \text{where } b_2 = \left[b_L; -b_U \right]$$

We denote the feasible set by $\mathcal{F} \equiv \{z_t \in \mathbb{R}^{4n} : Cz_t - d \ge 0\}.$

The two-step gradient projection method [8] defines its iterates $z_t^{(k+1)}$ from the previous iterate $z_t^{(k)}$ by first projecting onto the feasible set the vector defined by a steepest descent method:

$$\bar{z}_t^{(k)}\left(\alpha^{(k)}\right) = P\left(z_t^{(k)} - \alpha^{(k)}\nabla\Phi\left(z_t^{(k)}\right)\right),\tag{5}$$

where P is the projection operator onto the feasible set and $\alpha^{(k)} > 0$ helps guarantee convergence. We then perform a linesearch along this direction to obtain a suitable steplength $v_t^{(k)}$:

$$z_t^{(k+1)} = z_t^{(k)} + v_t^{(k)} \left(\bar{z}_t^{(k)} \left(\alpha^{(k)} \right) - z_t^{(k)} \right).$$

Our estimates to the coefficients are now defined as $\theta_t^{(k+1)} = \widetilde{I_1} z_t^{(k+1)}$ and $\theta_{t+1}^{(0)} = \widetilde{I_2} z_t^{(k+1)}$.

For ease of notation, we drop the superscripts corresponding to the iterates by denoting the current iterate z_t . We define

$$z_t(\alpha) = z_t - \alpha \nabla \Phi(z_t) \text{ and } \bar{z}_t(\alpha) = P(z_t - \alpha \nabla \Phi(z_t)).$$

Iterates defined in this manner guarantee the objective function will decrease at each iteration: if z_t is a feasible point and if \bar{z} is not a stationary point, then there exists a scalar $\bar{\alpha} > 0$ such that $\Phi(\bar{z}_t(\alpha)) < \Phi(z_t)$ for all $\alpha \in (0, \bar{\alpha}]$; and z_t is stationary if and only if $\bar{z}_t(\alpha) = z_t$ for all $\alpha \ge 0$ (see [8]).

3.2. Projecting Onto the Feasible Set

The projection of $\bar{z}(\alpha)$ onto the feasible set \mathcal{F} is the closest point $\bar{z}_{P}(\alpha)$ in the feasible set in Euclidean norm, i.e.,

$$\bar{z}_t(\alpha) \equiv \underset{z \in \mathbb{R}^{4n}}{\operatorname{argmin}} \quad \pi(z) = \frac{1}{2} \|z - z_t(\alpha)\|_2^2 \qquad (6)$$

subject to $Cz - d \ge 0$

The constraints on this minimization problem make it difficult to solve. However, we can simplify this problem by considering its *dual*. Specifically, the function $\mathcal{L}(z,\mu) = \frac{1}{2} ||z - z_t(\alpha)||_2^2 - \mu^T (Cz_t - d)$ is the Lagrangian function $\mathcal{L} : \mathbb{R}^{4nx8n} \to \mathbb{R}$ with Lagrange multipliers $\mu \in \mathbb{R}^{8n}$. The Lagrange dual function $g : \mathbb{R}^{8n} \to \mathbb{R}$ is given by $g(\mu) = \inf_{z \in \mathcal{F}} \mathcal{L}(z,\mu)$. By looking at $\nabla_z \mathcal{L}(z,\mu) = 0$, we find that

$$z = z_t \left(\alpha \right) + C^T \mu. \tag{7}$$

The dual associated with (6), which is subject to $\mu \ge 0$, is :

$$\mu^{*} \equiv \operatorname*{argmax}_{\mu \in \mathbb{R}^{8n}} g(\mu) = -\frac{1}{2} \left\| C^{T} \mu \right\|_{2}^{2} - \mu^{T} \left(C \bar{z}(\alpha) - d \right)$$
(8)

3.3. Duality Gap

In the primal problem (6), we note that the objective function $\pi(z)$ is convex and the constraints $Cz_t \ge d$ are affine. If we let $b = \frac{1}{2}(b_U + b_L)$, $u = [W^{-1}b]_+$, $v = [-W^{-1}b]_+$, then the feasible set is non-empty. A weaker version of Slater's condition (see [9], p.226) has therefore been satisfied, implying that the duality gap is zero, i.e. $g(\mu^*) = \pi(\bar{z}_t(\alpha))$. The solution to (6) can thus be defined as $\bar{z}_t(\alpha) \equiv z_t(\alpha) + C^T \mu^*$ by solving (8).

3.4. Solving the Dual Problem

We partition μ as $\mu = [\lambda; \zeta; \psi]$ where $\lambda \in \mathbb{R}^{4n}$, and $\zeta, \psi \in \mathbb{R}^{2n}$. The Lagrange multipliers λ, ζ , and ψ correspond to the constraints $z_t \ge 0, b_U \ge W\theta_t \ge b_L$, and $b_U \ge W(\theta_t + \Delta \theta_t) \ge b_L$ respectively. The dual problem can now be equivalently written as :

It is worth noting that (9) has only bound constraints while (6) has linear constraints which are generally more difficult to satisfy. We can solve (9) using component-wise minimization, where each component is obtained sequentially by treating the other components as constant and taking the partial derivatives of $h(\lambda, \zeta, \psi)$ and setting them equal to 0. We let $\zeta = [\zeta_1; \zeta_2], \psi = [\psi_1; \psi_2]$ where $\zeta_1, \zeta_2, \psi_1, \psi_2 \in \mathbb{R}^n$. Using the steps outlined in [4], which examines the first-order optimality conditions, we can show

$$\begin{split} \lambda^{(j)} &= \left[-\left(\left[\begin{array}{cc} \widetilde{W_1}^T & -\widetilde{W_1}^T \end{array} \right] \zeta^{(j-1)} \right. \\ &+ \left[\begin{array}{cc} \widetilde{W_2}^T & -\widetilde{W_2}^T \end{array} \right] \psi^{(j-1)} + z_t \left(\alpha \right) \right) \right]_+ \\ \zeta_1^{(j)} &= \left. \frac{1}{2} \left[-r_L^j \right]_+ , \qquad \zeta_2^{(j)} = \frac{1}{2} \left[-r_U^j \right]_+ \\ \psi_1^{(j)} &= \left. \frac{1}{4} \left[-s_L^j \right]_+ , \qquad \psi_2^{(j)} = \frac{1}{4} \left[-s_U^j \right]_+ \end{split}$$

where

$$\begin{split} r_L^j &= + \left[2 \left(\psi_1^{(\widetilde{j})} - \psi_2^{(\widetilde{j})} \right) + \widetilde{W}_1 \left(\lambda^{(j)} + z_t \left(\alpha \right) \right) \right] - b_L \\ r_U^j &= - \left[2 \left(\psi_1^{(\widetilde{j})} - \psi_2^{(\widetilde{j})} \right) + \widetilde{W}_1 \left(\lambda^{(j)} + z_t \left(\alpha \right) \right) \right] + b_U \\ s_L^j &= + \left[2 \left(\zeta_1^{(j)} - \zeta_2^{(j)} \right) + \widetilde{W}_2 \left(\lambda^{(j)} + z_t \left(\alpha \right) \right) \right] - b_L \\ s_U^j &= - \left[2 \left(\zeta_1^{(j)} - \zeta_2^{(j)} \right) + \widetilde{W}_2 \left(\lambda^{(j)} + z_t \left(\alpha \right) \right) \right] + b_U \end{split}$$

3.5. Feasibility

At each dual problem iteration, an estimate for the primal problem iterates, $\theta_t^{(k)}$ and $\theta_{t+1}^{(k)}$, can be shown to be feasible. If we let $\mu^{(j-\frac{1}{3})} = (\lambda^{(j)}, \zeta^{(j)}, \psi^{(j-1)})$ and $z_t^{(j-\frac{1}{3})} = z_t(\alpha) + C^T \mu^{(j-\frac{1}{3})}$ from (7), then during the $j - \frac{1}{3}$ iteration of the component-wise minimization, the primal variable associated with the dual variable is feasible with respect to the constraints in (5):

$$W\theta_{t}^{(j)} = \widetilde{W}_{2}z_{t}^{(j-\frac{1}{3})} = \widetilde{W}_{2}\left(z_{t}(\alpha) + C^{T}\mu^{(j-\frac{1}{3})}\right)$$

$$= \left[\widetilde{W}_{2}\left(z_{t}(\alpha) + \lambda^{(j)}\right) + 2\left(\psi_{1}^{(j)} - \psi_{2}^{(j)}\right)\right]$$

$$+ 2\left(\zeta_{1}^{(j)} - \zeta_{2}^{(j)}\right)$$

$$= b_{L} + \left[r_{L}^{j}\right]_{+} - \left[-r_{U}^{j}\right]_{+}$$

We can also show that $W\theta_t^{(j)} = b_U + [-r_L^j]_+ - [r_U^j]_+$. Using these equations, we can show that $b_U \ge W\theta_t^{(j)} \ge b_L$, i.e., the

iterates generated by this component-wise minimization approach are feasible.

Additionally, at the end of the *j*-th iteration of the componentwise minimization, we obtain a feasible estimate for initial value $\theta_{t+1}^{(0)}$ for the next optimization problem by the following: the primal variable associated with the dual variable $\mu^{(j)} = (\lambda^{(j)}, \zeta^{(j)}, \psi^{(j)})$, is given by $z_t^{(j)} = z_t(\alpha) + C^T \mu^{(j)}$. By considering

$$W\theta_{t+1}^{(0)} = \left[\widetilde{W}_{1} \left(z_{t} \left(\alpha \right) + \lambda^{(j)} \right) + 2 \left(\zeta_{1}^{(j)} - \zeta_{2}^{(j)} \right) \right] \\ + 4 \left(\psi_{1}^{(j)} - \psi_{2}^{(j)} \right).$$

we can show that $b_U \ge W \theta_{t+1}^{(0)} \ge b_L$. Thus, by appropriately defining $(\theta_t^{(j)}, \theta_{t+1}^{(0)})$ within the above sequence of steps we can show that our approximate solution will maintain feasibility at every iteration.

Finally, we note that the order in which the components $\lambda^{(j)}$, $\gamma^{(j)}$, and $\psi^{(j)}$ of the Lagrange multipliers are computed is not unique. Other permutations will also give feasible primal iterates and will result in decreasing duality gaps. In our experience, however, the ordering we described in this section achieves smaller duality gaps in fewer iterations.

4. NUMERICAL RESULTS

We compare the results of the LCGPk methods from [5] with nonnegative constraints to our proposed Linearly Constrained Gradient Projection with arbitrary lower and upper bounds (LCGPk-UL) to examine the effectiveness of our method, where k is the number of frames solved simultaneously. We consider a video of a signature (in this case, of one of the authors) written electronically, which occurs in many credit card and identity verification applications. The video consists of 50 frames of 128 x 256 gray-scale frames (see Fig. 1(a)). We can reconstruct the video whose frames have been blurred by noise (whether it be from faulty equipment or data corruption) with an accurate approximation of the original signature. The observations are created by passing the true video ${f_t^*}_{t=1}^{50}$ through the optical blurring operation A_t using the 2D blur kernel $h_{i,j} = 1/(1+i^2+j^2)$, for i, j = -4, ..., 4, which is the same kernel used in [7]. We constrain the solution to lie between 0 and 255 which are the pixel intensity constraints. This video contains regions of both significant darkness and brightness, making it a useful test for our method. We have optimized the τ values to yield the minimum root-mean-squared error $||f - \hat{f}||_2 / ||f||_2$ for the last 30 frames, which is when the RMS of both methods have approximately reached a steady state. Table 1 shows the RMS for the 50th frame for LCGPk and LCGPk-UL, where k frames is solved simultaneously with 3 seconds allotted per frame.

Method	RMS(%)	Method	RMS(%)
LCGP2	3.681	LCGP2-UL	2.407
LCGP4	3.390	LCGP4-UL	2.908

Table 1. Reconstruction RMS for single-trial results and results averaged over ten trials. $RMS(f) \equiv ||f - \hat{f}||_2/||f||_2$. We note that the LCGP2 and LCGP4 approaches only impose lower bound constraints, whereas the LCGP2-UL and LCGP4-UL approaches impose both lower and upper bound constraints.

We notice that the LCGPk-UL method has a significantly lower RMS than its LCGPk counterpart. In fact, LCGP2-UL has a lower



Fig. 1. Results of our numerical experiments. Here we show (a) True intensity f^* at 50th frame (b) the blurred observation y (c) the reconstruction using LCGP2 (RMS = 3.681%), (d) the reconstruction with the proposed LCGP2-UL method (RMS = 2.407%) (e) RMS evolution at each time frame.

RMS than LCGP2 on every frame, while LCGP4-UL shows a lower RMS on 88% of the frames compared to LCGP4. Examining Fig.1(c) reveals the weakness of having no upper constraint; intensity values rise above 340, far exceeding the natural limit of 255. The scale in Fig.1(c) is different from the other figures to emphasize (1) the significant overestimation of pixel intensity values when an upper bound constraint is not imposed, and (2) the ringing artifacts present in this reconstruction. The upper constraints of the LCGPk-UL method help contain that, as seen in Fig.1(d). Increasing the number of frames solved simultaneously but keeping the time allotted constant causes each algorithm to perform fewer reconstruction iterations, making the initial frames less accurate. We note that the cursor '+' at the tail end of the signature appears blurry in the reconstruction. This is due to the fact that the cursor is constantly moving in addition to the being blurry, which makes it challenging to recover. In [5], it was noted that k = 4 was the cutoff in increased accuracy; higher values of k resulted in insufficient iterations in the allotted time. In our work, the optimal number of frams to solve simultaneously is k = 2, as can be observed in Table 1. Thus, we have demonstrated that the general bound constrained multi-frame method significantly outperforms the non-negative constrained method.

5. CONCLUSION

This paper presents the gradient projection method LCGPk-UL for video reconstruction that simultaneously solves for multiple frames. The accuracy of our reconstruction increases when we take advantage of the correlation between frames and enforce both the sparsity constraints and the upper/lower bound constraints (as opposed to just non-negative constraints). We have demonstrated through a numerical experiment that the proposed LCGPk-UL can improve upon the performance of currently available gradient projection methods. The results also suggest that larger multi-frame methods can yield reconstructions with increased accuracy by increasing the allotted time, which involves only slightly more computational effort.

6. REFERENCES

- E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 15, no. 12, pp. 4203–4215, 2005.
- [2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] R. Tibshirani, "Regression shrinkage and selection via the lasso," J. Roy. Statist. Soc. Ser. B, vol. 58, no. 1, pp. 267–288, 1996.
- [4] J. Hernandez, D. Thompson, Z. Harmany, and R.F. Marcia, "Bounded gradient projection methods for sparse signal recovery," in *Proc. 2011 IEEE Int. Conf. on Acoustic, Speech, and Signal Process.*, Prague, Czech Republic, 2011.
- [5] D. Thompson, Z. Harmany, and R.F. Marcia, "Sparse video recovery using linearly constrained gradient projection," in *Proc.* 2011 IEEE Int. Conf. on Acoustic, Speech, and Signal Process., Prague, Czech Republic, 2011.
- [6] R.F. Marcia and R.M. Willett, "Compressive coded aperture video reconstruction," in *Proc. 16th European Signal Process. Conf.*, Lausanne, Switzerland, August 2008.
- [7] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Sign. Proces.: Special Issue on Convex Optimization Methods for Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.
- [8] P. H. Calamai and J. J. More, "Projected gradient methods for linearly constrained problems," *Math. Programming*, vol. 39, no. 1, pp. 93–116, 1987.
- [9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.