CARTOON-LIKE IMAGE RECONSTRUCTION VIA CONSTRAINED $\ell_p$-MINIMIZATION

Simon Hawe, Martin Kleinsteuber, and Klaus Diepold

Department of Electrical Engineering and Information Technology, Technische Universität München, München, Germany.
e-mail: \{simon.hawe,kleinsteuber,kldi\}@tum.de.

Abstract
This paper considers the problem of reconstructing images from only a few measurements. A method is proposed that is based on the theory of Compressive Sensing. We introduce a new prior that combines an $\ell_p$-pseudo-norm approximation of the image gradient and the bounded range of the original signal. Ultimately, this leads to a reconstruction algorithm that works particularly well for Cartoon-like images that commonly occur in medical imagery. The arising optimization task is solved by a Conjugate Gradient method that is capable of dealing with large scale problems and easily adapts to extensions of the prior. To overcome the none differentiability of the $\ell_p$-pseudo-norm we employ a Huber-loss term like approximation together with a continuation of the smoothing parameter. Numerical results and a comparison with the state-of-the-art methods show the effectiveness of the proposed algorithm.

Index Terms— Compressive Sensing, $\ell_p$ minimization, Image Reconstruction, Conjugate Gradient Algorithm

1. INTRODUCTION AND NOTATION

In recent years, Compressive Sensing (CS) \cite{1, 2} has evolved as one of the most active research topics in the signal processing community. Basically, CS is a joint sampling and compression mechanism, which enables perfect signal reconstruction from a very small number of non-adaptively acquired measurements. These measurements encode the entire information about the signal at hand into a very small amount of data, much smaller than the signal’s dimension. More precisely, let $s \in \mathbb{R}^n$ denote a set of signal independent measurement vectors $\{\phi_i\}_{i=1}^m$, which can be compactly written as

$$y = [\phi_1, \ldots, \phi_m]^T s + e =: \Phi s + e. \tag{1}$$

The vector $y \in \mathbb{R}^m$ contains the measurements, $e \in \mathbb{R}^m$ models sampling errors, and $\Phi \in \mathbb{R}^{m \times n}$ is the measurement matrix. Clearly, without further information on the signal, the problem of inferring the signal from the measurements is ill-posed. Prior assumptions on the signal help to well-define the recovery problem: The CS-framework exploits the fact that many interesting signals have a sparse- or compressible representation with respect to some (possibly overcomplete) basis.

Let $x \in \mathbb{R}^d$ with $d \geq n$ denote the $k$-sparse representation of the signal, where $k$-sparse means that only $k \ll d$ entries of $x$ are nonzero. We write the corresponding linear transformation as $x = Ds$, where $D \in \mathbb{R}^{d \times n}$ is a sparsifying transformation with full rank. Furthermore, let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function that promotes or measures sparsity. We denote the Moore-Penrose pseudo-inverse of $D$ by $D^\dagger$. The recovery of the signal thus leads to the well known synthesis approach, cf. \cite{3},

$$\min_{x \in \mathbb{R}^d} g(x) \quad \text{subject to} \quad \|\Phi D^\dagger x - y\|_2^2 \leq \epsilon. \tag{2}$$

where $\epsilon$ is an estimated upper bound on the noise power $\|e\|_2^2$. Informally speaking, by solving \eqref{eq:2}, we find the sparsest vector $\hat{x}$ that is compatible with the acquired measurements. The signal is then recovered by $\hat{s} = D\hat{x}$.

Another common procedure is to directly search for a signal $\hat{s}$ such that $D\hat{s}$ is sparse. Formally, $\hat{s}$ is the solution of

$$\min_{s \in \mathbb{R}^n} g(Ds) \quad \text{subject to} \quad \|\Phi s - y\|_2^2 \leq \epsilon. \tag{3}$$

which is known as the analysis approach, see \cite{3} for the relation between problem \eqref{eq:2} and \eqref{eq:3}. In the field of image processing, problem \eqref{eq:3} is favored over \eqref{eq:2} due to the lower dimension of the search space. Moreover, for many practically important operators $D$ the computation of $D^\dagger x$ is infeasible. Thus, we will focus on problem \eqref{eq:3} here. Throughout the paper, $v(i)$ denotes the $i^{th}$ entry of the vector $v$ and $M(i,j)$ the $(i,j)$-entry of the matrix $M$.

Clearly, with the $k$-sparsity assumption on $x$, the ideal choice for $g$ would be the $\ell_0$-pseudo-norm $\|v\|_0 := \#\{i : v(i) \neq 0\}$, which counts the nonzero elements of $v$. Unfortunately, solving problem \eqref{eq:3} with $g(Ds) = \|Ds\|_0$ is computationally intractable as it is combinatorial NP-hard. Instead, it has been shown in \cite{4} that depending on $\Phi$ and $D$...
the replacement of the $\ell_0$-pseudo-norm by its closest convex
surrogate, the $\ell_1$-norm $\|v\|_1 := \sum_i |v(i)|$, gives the same
solution. In the noise free case, if the number of measure-
ments $m$ is large enough compared to the sparsity factor $k$,
this solution to (3) yields the exact recovery of the signal
[5]. The number of measurements required for ideal recon-
struction can even be further decreased, if an $\ell_p$-pseudo-norm
$\|v\|_p := \sum_i |v(i)|^p$ with $0 < p < 1$ is employed instead of the
$\ell_1$-norm, [6]. However, the resulting optimization prob-
lem cannot be solved by straightforward linear quadratic pro-
gramming, but approximations like reweighted least squares
[7], or an iterative shrinkage method [8] can applied.

In this paper we present an algorithm for reconstructing
compressively sampled images via $\ell_p$-minimization. We
introduce an additional prior that copes with the boundedness of
the range of images. This prior accelerates the optimization
process and drastically improves the reconstruction results.

2. PROBLEM STATEMENT

As stated above, our goal is to reconstruct compressively sam-
ples images $I \in \mathbb{R}^{h \times w}$, where our focus is on Cartoon-like
images, i.e. images that are piecewise flat.

We consider two sampling bases, namely partial discrete
Fourier transform (DFT) and the Rudin-Shapiro transfor-
mation (RST) [9]. The interest for the DFT arises due to its im-
portance in magnetic resonance imaging (MRI) where Fourier
coefficients are directly sampled. The RST, also know as the
real valued Dragon-Noiselet-transformation, is used as it can be
implemented into a real imaging sensor, and due to its de-
sirable properties for image reconstruction [10].

Following a common approach in image reconstruction,
we impose a sparsity prior on the image gradient. The sim-
plest way of approximating the gradient is in terms of finite
differences with

$$\frac{\partial I}{\partial x}(i, j) = \begin{cases} (I(i, j) - I(i, j+1))/\sqrt{2} & \text{if } i < w \\ 0 & \text{otherwise,} \end{cases}$$

being the difference between two neighboring pixels in hor-
izontal direction, and $\frac{\partial I}{\partial y}(i, j)$ the difference between two
neighboring pixels in vertical direction defined accordingly.
Let $s := \text{vec}(I) \in \mathbb{R}^n$ with $n = hw$ be the vectorized
image obtained by stacking its columns among each other. We
define $D_x \in \mathbb{R}^{n \times n}$ and $D_y \in \mathbb{R}^{n \times n}$ as those matrices that
realize the approximate image gradients, i.e.

$$D_x s = \text{vec} \left( \frac{\partial I}{\partial x} \right), \quad D_y s = \text{vec} \left( \frac{\partial I}{\partial y} \right).$$

The sparsity assumption on the image gradient yields

$$D = [D_x \quad D_y] \in \mathbb{R}^{2n \times n},$$

as the sparsifying transformation that we consider for our
problem at hand.

Regarding the choice of the sparsity measure, we com-
pared the $p$-pseudo-norms $\|Ds\|_p$ with

$$\|Ds\|_{TV} := \sqrt{\sum_{i=1}^n ((Ds)(i)^2 + (Ds)(i+n)^2)^{\frac{q}{2}}}$$

which for $p = 1$ are the well known anisotropic, and isotropic
Total Variation pseudo-norms, respectively. While the $p$-
pseudo-norm enforces sparsity of the gradient in $x$ and $y$
direction separably, the latter (7) enforces the magnitude
of the gradient to be sparse. This has the nice effect that (7)
is invariant under rotations of the underlying picture, conse-
quently, we employ it for further studies here.

Finally, we exploit the fact that the pixel intensities of im-
gages are bounded. If we denote the lower bound by $\tau_l$ and
the upper bound by $\tau_u$, this leads to the constraint optimization
problem

$$\text{minimize } \|Ds\|_{TV},$$

subject to $\|\Phi s - y\|_2^2 \leq \epsilon, \tau_l \leq s(i) \leq \tau_u.$

For $\tau_l = 0$ and $\tau_u = +\infty$ this is the well known positivity
constraint. Here, as we are dealing with images, we typically
chose $\tau_l = 0$ and $\tau_u = 1$ or $\tau_u = 255$ depending on the
image format. Our experiments provide evidence that these
additional constraints drastically improve the reconstruction
quality and accelerate the optimization process.

3. RECONSTRUCTION ALGORITHM

The reconstruction algorithm proposed here is based on a
Conjugate Gradient (CG) method that minimizes a smooth
approximation of the optimization problem (8). It allows
to easily integrate additional priors into the minimization
process without sever changes. To enhance legibility, we
stick to the matrix-vector notation. However, note that all
matrix-vector-products are efficiently implemented via filter-
ing techniques in at least $O(n \log n)$ flops.

The approximation is carried out in three ways. First, the
bounding constraint is enforced by the functional $\tau(s) :=
\sum_{i=1}^n \tau(s(i))$ where $\tau$, is a penalty term defined as

$$\tau(x) = \begin{cases} |x - \tau_u|^q & \text{if } x \geq \tau_u \\ |x - \tau_l|^q & \text{if } x \leq \tau_l \\ 0 & \text{otherwise,} \end{cases},$$

with $q > 1$. The larger $q$, the higher the penalty and the tighter
the bound. For the presented image reconstruction algorithm
$q = 2$ yields very good fit to the range.

Second, since the CG algorithm requires a differentiable
cost function, we employ a smooth approximation of (7) that
extends the well known Huber loss term to the $\ell_p$-pseudo-
norm. We define the $p$-Huber loss as

$$h_{p,\mu}(x) = \begin{cases} |x|^p - \kappa_1 & \text{if } |x| \geq \mu \\ \kappa_2 x^2 & \text{otherwise,} \end{cases}$$
with $\kappa_1 = (1 - \frac{p}{2})\mu^p$ and $\kappa_2 = \frac{p}{2}\mu^{p-2}$. Note, that $h_{p,\mu}(x)$ is differentiable for all $p > 0$ and $\mu > 0$. Combining (7) and (10) leads to the following approximation of the sparsity measure for the image gradient

$$\|D\|_{TVp} \approx \sum_{i=1}^{n} h_{\mu,p} \left( \sqrt{(D_s(i))^2 + (D_s(i+n))^2} \right)$$

$$= \|D\|_{\text{TVp,}\mu}. \quad (11)$$

Ultimately, problem (8) is approximated in unconstrained Lagrangian form as

$$\min_{s \in \mathbb{R}^n} f(s) = \frac{1}{2} \|\Phi s - y\|^2 + \tau(s) + \lambda \|D\|_{TVp,\mu}. \quad (12)$$

The Lagrange multiplier $\lambda \in \mathbb{R}_0^+$ weighs between the sparsity of the solution and its fidelity to the acquired samples according to $\lambda \sim \epsilon$.

The CG method for minimizing (12) is initiated with $s_0 = \Phi^* y$ and iteratively updates the current solution by

$$s_{i+1} = s_i + \alpha_i h_i. \quad (13)$$

The scalar $\alpha_i \geq 0$ is the line-search parameter or the stepsize, and $h_i$ is the descent direction at the $i^{th}$ iteration. Various line-search techniques for finding $\alpha_i$ that approximately solve

$$\min_{\alpha \in \mathbb{R}} f(s_i + \alpha h_i) \quad (14)$$

exist from which we choose backtracking line-search [11] as it is conceptually simple and computationally cheap.

Let $g_i := \nabla f(s_i)$ be the gradient of the cost function (12) at the $i^{th}$ iteration and let the descent direction be initiated with $h_0 = -g_0$. The authors showed in [12] that the Hestenes-Stiefel formula

$$h_{i+1} = -g_{i+1} + \frac{g_{i+1} - g_i}{\nabla f(s_i)} h_i. \quad (15)$$

is well suited for image reconstruction. The algorithm updates $s_i$ and $h_i$ as in (13) and (15) until the subsequently defined stopping criterion is met.

Following [13], the algorithm terminates if the relative variation of the regularizing function over the last $l$ iterations

$$\gamma = (\|\tilde{g}(D s_i) - \tilde{g}_i\|/\tilde{g}_i) \quad (16)$$

with $\tilde{g}_i = \frac{1}{l} \sum_{k=1}^{l} g(D s_i - k)$, falls below a certain threshold $\delta$. Typically, we choose $l \in [10, 20]$ and $\delta \in [10^{-6}, 10^{-10}]$, depending on the required accuracy. The CG-method is summarized in Algorithm 1.

Regarding the smoothing parameter $\mu$ and the Lagrange multiplier $\lambda$, large values lead to fast convergence but yield biased reconstruction results. To overcome this problem, the reconstruction algorithm consists of repeating Algorithm 1 for a predetermined number of times, say $N$. It is initiated with relatively large $\mu_0$ and $\lambda_0$ and updates $\lambda_{k+1} = c_\lambda \lambda_k$, $\mu_{k+1} = c_\mu \mu_k$ such that $\lambda_N$ and $\lambda_N$ take a predefined value $\mu_f \ll \mu_0$ and $\lambda_f \ll \lambda_0$. To that end, the continuation parameters are chosen as

$$c_\lambda := \left( \frac{\lambda_f}{\lambda_0} \right)^{1/N}, c_\mu := \left( \frac{\mu_f}{\mu_0} \right)^{1/N}. \quad (17)$$

The complete reconstruction algorithm is summarized in Algorithm 2.

Algorithm 1 CG Algorithm

Input: $y, s_0, p, \lambda, \mu, \delta$
Set: $g_0 := \nabla f(s_0)$
$h_0 := -g_0$
for $i = 0$ to $\#\text{InnerIter}$ do
\begin{itemize}
  \item compute stepsize $\alpha_i$ via backtracking linesearch
  \item $s_{i+1} := s_i + \alpha_i h_i$
\end{itemize}
If $\gamma < \delta$, then stop (cf. (16))
Update search direction $h_{i+1}$ as in (15)
end for
Output: $s_i$

Algorithm 2 Image Reconstruction via CG

Input: $y, p, \lambda_0, \mu_0, \lambda_f, \mu_f, \delta, N$
Set: $s_0 := \Phi^* y$, compute $c_\lambda, c_\mu$ as in (17)
for $k = 0$ to $N$ do
\begin{itemize}
  \item $s^* := \text{CG-method}(y, s_0, p, \lambda_k, \mu_k, \delta)$
  \item $\lambda_{k+1} := c_\lambda \lambda_k$, $\mu_{k+1} := c_\mu \mu_k$
  \item $s_0 := s^*$
\end{itemize}
end for
Output: $s^*$

4. RESULTS

In this section we present some numerical experiments that reveal the performance of our algorithm. The test images are the Shepp-Logan Phantom ($I_1$) from [2], a MRI brain image ($I_2$) Figure 1(a), a cartoon image ($I_3$) Figure 1(d), and the famous Cameraman ($I_4$). The intensities of the images have been scaled to [0, 255].

For $I_1$ and $I_2$, DFT is used for sampling, whereas RST is used for $I_3$ and $I_4$. We measure the reconstruction quality in terms of the relative reconstruction error $\text{Rel} = \|s - s^*\|_2/\|s\|_2$ and in terms of PSNR $= 20 \log_{10} \left( \frac{255n}{\|s - s^*\|_2^2} \right) dB$.

Our method is compared with the NESTA algorithm [13]. Certainly, the same set of samples and the same stopping criterion is used for all algorithms.

To illustrate the influence of $p$, we present the results of our method with $p = 0.9$, $p = 0.7$, and $p = 0.4$. It can be clearly seen in Table 1 that the proposed CG method outperforms NESTA in all situations for the chosen test images, in particular for the noiseless Cartoon-like images $I_1$ and $I_3$ but
also for the natural image $I_4$. We like to mention two further results from the literature on the reconstruction of $I_1$ from a few Fourier samples employing an $\ell_p$ like sparsity measure. In [7] the authors report a relative reconstruction error of $\approx 2e^{-3}$ with $m/n = 0.038$ and in [8] an error of $\approx 6.68e^{-10}$ with $m/n = 0.035$. Note that our method yields a reconstruction error of $\approx 8.31e^{-9}$ with only $m/n = 0.025$, i.e. 30% less samples. Finally, Figures 1(b)-(e) present some reconstructed images. It can be seen that our method yields less blurry and smeared images than NESTA. The cartoon 1(d) is quasi perfectly reconstructed.

From numerous experiments with various images we conclude that $p = 0.7$ is a universally good choice. Very good reconstruction results have been observed after $N = 3$ iterations of Algorithm 1. Considering the smoothing parameter $\mu$, an initial value of $\mu_0 = 0.2$ and final value of $\mu_f = 1e^{-10}$ is well suited for general images. The value of $\lambda$ depends on the assumption about the noise level in the image. In the noise free case, we recommend $\lambda_k = \mu_k$.

5. CONCLUSION

In this paper we presented an algorithm for reconstructing images. The proposed method relies on the theory of compressive sampling and requires only a few samples compared to the image resolution. It is particularly well suited for Cartoon-like images, which frequently occur in medical imagery. An adaption to any sparsity measure based on the $\ell_p$-pseudonorm, to other sparsifying transformations and also to other signals is possible in a straightforward manner. A Matlab code is available on the webpage www.gol.ei.tum.de.

6. REFERENCES


Table 1: Reconstruction quality in PSNR (in dB) and the relative reconstruction error (Rel); $m$: number of samples; $n$: signal dimension.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Phantom, $I_1$</th>
<th>Brain, $I_2$</th>
<th>Cartoon, $I_3$</th>
<th>Cameraman, $I_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>m/n = 0.025</td>
<td>PSNR Rel</td>
<td>PSNR Rel</td>
<td>PSNR Rel</td>
<td>PSNR Rel</td>
</tr>
<tr>
<td>CG $p = 0.9$</td>
<td>31.5</td>
<td>3.14e$^{-1}$</td>
<td>30.0</td>
<td>1.44e$^{-1}$</td>
</tr>
<tr>
<td>CG $p = 0.7$</td>
<td>31.5</td>
<td>6.06e$^{-1}$</td>
<td>31.1</td>
<td>1.27e$^{-1}$</td>
</tr>
<tr>
<td>CG $p = 0.4$</td>
<td>173.7</td>
<td>8.31e$^{-1}$</td>
<td>30.8</td>
<td>1.32e$^{-1}$</td>
</tr>
<tr>
<td>NESTA</td>
<td>24.7</td>
<td>2.43e$^{-1}$</td>
<td>29.9</td>
<td>1.66e$^{-1}$</td>
</tr>
</tbody>
</table>

Fig. 1: Reconstruction results for $I_2$ and $I_3$