SINGLE-FRAME SIGNAL RECOVERY USING A SIMILARITY-PRIOR BASED ON PEARSON TYPE VII MRF

Sakinah Ali Pitchay and Ata Kabán
School of Computer Science, University of Birmingham, Edgbaston, Birmingham, B15 2TT, U.K.

Keywords: Single frame super-resolution, Compressive sensing, Similarity prior, Image recovery.

Abstract: We consider the problem of signal reconstruction from noisy observations in a highly under-determined problem setting. Most of previous work does not consider any specific extra information to recover the signal. Here we address this problem by exploiting the similarity between the signal of interest and a consecutive motionless frame. We incorporate this additional information of similarity that is available into a probabilistic image prior based on the Pearson type VII Markov Random Field model. Results on both synthetic and real data of MRI images demonstrate the effectiveness of our method in both compressed setting and classical super-resolution experiments.

1 INTRODUCTION

Conventional image super-resolution (SR) aims to recover a high resolution scene from a single or multiple frames of low resolution measurements. A noisy frame of a single low resolution image or signal often suffers from a blur and down-sampling transformation. The problem is more challenging when the observed data is a single low resolution frame because it contains fewer measurements than the number of unknown pixels of the high resolution scene that we aim to recover. This makes the problem ill-posed and under-determined too. For this reason, some additional prior knowledge is vital to obtain a satisfactory solution. We have demonstrated in previous work (A. Kabán and S. AliPitchay, 2011) that the Pearson type VII density integrated with Markov Random Fields (MRF) is an appropriate approach for this purpose.

In this paper, we tackle the problem using a more specific prior information, namely the similarity to a motionless consecutive frame as the additional input for recovering the signals of interest in a highly under-determined setting. This has real applications e.g. in medical imaging where such frames are obtained from several scans. Previous work in (N. Vaswani and W. Lu, 2010) found the average frame from those scans to be useful for recovery.

In principle, the more information we have about the recovered signal, the better the recovery algorithm is expected to perform. This hypothesis seems to work in (JCR. Giraldo et al., 2010; N. Vaswani and W. Lu, 2010), however both of these works require us to tune the free parameters of the model manually, and (JCR. Giraldo et al., 2010) reckons that the range of parameter values was not exhaustively tested. (N. Vaswani and W. Lu, 2010) also mentions that they were not able to attain exact reconstruction using fewer measurements than those needed by compressed sensing (CS) for a small image. By contrary, in this paper we will demonstrate good recovery from very few measurements using a probabilistic model that includes an automated estimation of its hyper-parameters.

Related work on sparse reconstruction gained tremendous interest recently and can be found in e.g. (R. G. Baraniuk et al., 2010; S. Ji et al., 2008; E. Candes et al., 2006; DL. Donoho, 2006). The sparser a signal is, in some basis, the fewer random measurements are sufficient for its recovery. However these works do not consider any specific extra information that could be used to accentuate the sparsity, which is our focus. Somewhat related, the recent work in (W. Lu and N. Vaswani, 2011) exploits partial erroneous information to recover small image sequences.

This paper is aimed at taking these ideas further through a more principled and more comprehensive treatment. We consider the case when the observed frame contains too few measurements, but an additional motionless consecutive scene in high resolutions is provided as an extra input. This assumption is often realistic in imaging applications. Our aim is to reduce the requirements on the number of mea-

Ali Pitchay S. and Kabán A. (2012).
SINGLE-FRAME SIGNAL RECOVERY USING A SIMILARITY-PRIOR BASED ON PEARSON TYPE VII MRF.
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, pages 123-133
DOI: 10.5220/0003791401230133

surements by exploiting the additional similarity information. To achieve this, we employ a probabilistic framework, which allows us to estimate all parameters of our model in an automated manner. We conduct extensive experiments that show that our approach not only bypasses the requirement of tuning free parameters but it is also superior to a cross validation method in terms of both accuracy and computation time.

2 IMAGE RECOVERY FRAMEWORK

2.1 Observation Model

A model is good if it explains the data. The following linear model has been used widely to express the degradation process from the high resolution signal $\mathbf{z}$ to a compressed or low resolution noisy signal $\mathbf{y}$ (L. C. Pickup et al., 2007; H. He and L. P. Kondi, 2004; H. He and L. P. Kondi, 2003; RC. Hardie and KJ. Barnard, 1997):

$$\mathbf{y} = \mathbf{W}\mathbf{z} + \eta \tag{1}$$

where the high resolution signal denoted by $\mathbf{z}$ is an N-dimensional column vector and $\mathbf{y}$ is an Mx1 matrix representing the noisy version of the signal, with $M < N$.

In classical super-resolution, the transformation matrix $\mathbf{W}$ typically consists of blur and down-sampling operators. In our study, we also utilise random Gaussian compressive matrices $\mathbf{W}$ with entries sampled independent and identically distributed (i.i.d) from a standard Gaussian. Finally, η is the additive noise, assumed to be Gaussian with zero-mean and variance, σ^2.

2.2 The Similarity Prior

The construction of a generic prior for images, the Pearson type VII MRF prior was presented in (A. Kabán and S. AliPitchay, 2011). It is based on the neighbourhood features $\mathbf{Dz}$ where $\mathbf{D}$ makes the signal sparse. In this paper, we aim to recover both 1D and 2D signals using the additional similarity information. We define the entries of $\mathbf{D}$, i.e d_{ij} as follows:

$$d_{ij} = \begin{cases} 1 & \text{if } i = j; \\ -1/\# & \text{if i and j are neighbours;} \\ 0 & \text{otherwise.} \end{cases}$$

where # denotes the number of cardinal neighbours and it is 4 for images and 2 for 1D signals.

In general, the idea is that the main characteristic of any natural image is a local-smoothness. This means that the intensities of neighbouring pixels tend to be very similar. Hence, $\mathbf{Dz}$ will be sparse. Therefore, here we propose an enhanced prior to exploit more information that leads to more sparseness. By employing the given additional information of the consecutive image or signal, we will employ the difference, $\mathbf{f}$ between the recovered image, $\mathbf{z}$ and the extra information denoted as $\mathbf{s}$. Obviously the more pixels $\mathbf{z}$ and $\mathbf{s}$ have in common, the more smooth their difference will be. Figure 1 shows a few examples of histograms of the neighbourhood features $\mathbf{Dz}$ from real images, where the sparsity is entirely the consequence of the local smoothness. Additionally, we also show the histograms of the new neighbourhood features $\mathbf{Df}$ that includes the additional similarity information. We see the latter is a lot sparser than the former.

Then we can formulate the i-th feature in a vector form, with the aid of the i-th row of this matrix (denoted $\mathbf{D}_i$) as the following:

$$f_i - \frac{1}{\#} \sum_{j \in \#\ \text{neighb}(i)} f_j \quad = \quad \sum_{j=1}^{N} d_{ij} f_j = \mathbf{D}_i \mathbf{f} \tag{2}$$

Since our task is to encode the sparse property of signals, therefore this feature is useful: The difference between a pixel of the difference image f and the average of its neighbours is close to zero, almost everywhere except an the edges of the dissimilarity areas.

Plugging this into the Pearson-MRF density, we have the following prior, that we refer to as a *similarity prior*:

$$Pr(\mathbf{z}) = \frac{1}{Z_{Pr(\lambda,\nu)}} \prod_{i=1}^{N} \{(\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2 + \lambda\}^{-\frac{1+\nu}{2}} \tag{3}$$

where $Z_{Pr(\lambda,\nu)} = \int dz \prod_{i=1}^{N} \{(\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2 + \lambda\}^{-\frac{1+\nu}{2}}$ is the partition function that makes the whole probability density function integrate to one, and this multivariate integral does not have an analytic form.

2.3 Pseudo-likelihood Approximation

As in previous work (A. Kabán and S. AliPitchay, 2011), we employ a pseudo-likelihood approximation to the partition function $Z_{p(\lambda,\nu)}$. Replacing the approximation using the extra information into (3), we obtain the following approximate image model:

$$Pr(\mathbf{z}|\lambda,\nu) \approx \prod_{i=1}^{N} \frac{\Gamma\left(\frac{1+\nu}{2}\right) \lambda^{\nu/2} \{(\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2 + \lambda\}^{-\frac{1+\nu}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{\pi}} \tag{4}$$

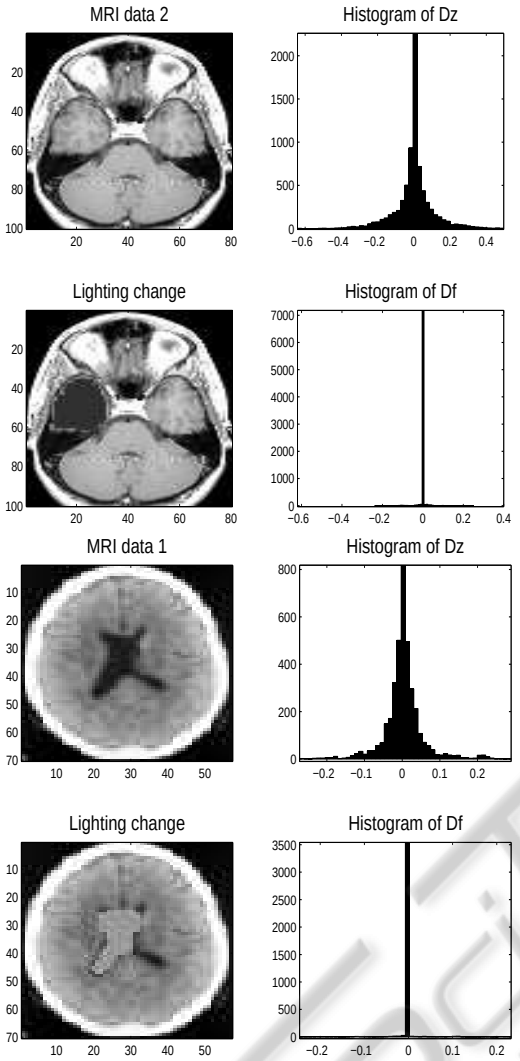


Figure 1: Example histograms of the distribution of neighbourhood features $\mathbf{D}_i\mathbf{z}$, and $\mathbf{D}_i\mathbf{f}$ where i=1,...,N from a MRI real data.

We shall employ this to infer $\mathbf{z}$ simultaneously with estimating our hyper-parameters λ, ν and σ.

2.4 Joint Model

The entire model is the joint model of the observations $\mathbf{y}$ and the unknowns $\mathbf{z}$.

$$Pr(\mathbf{y},\mathbf{z},f|\mathbf{W},\sigma^2,\lambda,\nu) = Pr(\mathbf{y}|\mathbf{z},\mathbf{W},\sigma^2)Pr(\mathbf{z}|f,\lambda,\nu) \quad (5)$$

where the first factor is the observation model and the second factor is the image prior model and its free parameters λ and ν.

3 MAP ESTIMATION

We will employ the joint probability (5) as the objective to be maximised. Maximising this w.r.t. $\mathbf{z}$ is also equivalent to finding the most probable image $\hat{\mathbf{z}}$, i.e. the maximum a posteriori (MAP) estimate, since (5) is proportional to the posterior $Pr(\mathbf{z}|\mathbf{y})$.

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}}\{-\log[Pr(\mathbf{y}|\mathbf{z})] - \log[Pr(\mathbf{z})]\} \quad (6)$$

Namely, the most probable high resolution signal is the one for which the negative log of the joint probability model takes its minimum value. Hence, our problem can be solved through minimisation. The expression for the negative log of the joint probability model will then be defined as our minimisation objective and also called as the error-objective. It can be written as:

$$Obj(\mathbf{z},\sigma^2,\lambda,\nu) = -\log[Pr(\mathbf{y}|\mathbf{z},\sigma^2)] - \log[Pr(\mathbf{z}|f,\lambda,\nu)] \quad (7)$$

Equation (7) may be decomposed into two terms: the first one that contains all the entries that involve $\mathbf{z}$ and the second one contains the terms that do not — i.e. Obj($\mathbf{z}$,σ^2,λ,ν)=$Obj_z(\mathbf{z})$ + $Obj_{(\lambda,\nu)}(\lambda,\nu)$.

3.1 Estimating the most Probable z

The observation model is also called the likelihood model because it expresses how likely it is that a given $\mathbf{z}$ produced the observed $\mathbf{y}$ through the transformation $\mathbf{W}$. Hence we have for the first term in (5):

$$Pr(\mathbf{y}|\mathbf{z}) \propto \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{W}\mathbf{z})^T(\mathbf{y}-\mathbf{W}\mathbf{z})\right\} \quad (8)$$

By plugging in the term for the observation model and the prior into (7), we obtain the objective function. The terms of the objective (7) that depend on $\mathbf{z}$ are the following:

$$Obj_z(\mathbf{z}) = \frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{W}\mathbf{z})^2 + \frac{\nu+1}{2}\sum_{i=1}^{N}\log\{(D_i(\mathbf{z}-\mathbf{s}))^2+\lambda\} \quad (9)$$

The most probable estimate is the $\hat{z}$ that has the highest probability in the model. It is equivalently the one that achieves the lowest error. Recap, our model has two factors which depend on the likelihood or also known as the observation model, and the image prior that assists the signal recovery. Thus, our error models both the *mismatch* of the predicted model $\mathbf{W}\mathbf{z}$ with the observed data $\mathbf{y}$ and *determinant* for allowing the free parameters to control the smoothness and the edges encoded in the prior.

The objective is differentiable; therefore any non-linear optimiser could be practical to optimise the term (9) w.r.t. $\mathbf{z}$. The gradient of the negative log likelihood term is given by:

$$\nabla(z)Obj_z = \frac{1}{\sigma^2}\mathbf{W}'(\mathbf{W}\mathbf{z}-\mathbf{y}) + (\nu+1)\sum_{i=1}^{N} D_i^T \frac{D_i(\mathbf{z}-\mathbf{s})}{(D_i(\mathbf{z}-\mathbf{s}))^2+\lambda} \quad (10)$$

3.2 Estimation of σ^2, λ and ν

Writing out the terms in (7) that depend on σ^2, we obtain a closed form for estimating the σ^2.

$$\sigma^2 = \frac{1}{M}\left(\sum_{i=1}^{M}(y_i - \mathbf{W}_i\mathbf{z})^2\right) \quad (11)$$

Terms that depend on λ and ν are given by:

$$Obj_{(\lambda,\nu)} = N\log\Gamma\left(\frac{1+\nu}{2}\right) - N\log\Gamma\left(\frac{\nu}{2}\right) + \frac{N\nu}{2}\log\lambda - \frac{1+\nu}{2}\sum_{i=1}^{N}\log((\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2+\lambda) \quad (12)$$

Both of these hyperparameters need to be positive valued. To ensure our estimates are actually positive, we parameterise the log probability objective (12) such as to optimise for the +/- square root of these parameters. Taking derivatives w.r.t $\sqrt{\lambda}$ and $\sqrt{\nu}$, we obtain:

$$\frac{d\log p(\mathbf{z})}{d\sqrt{\lambda}} = \sum_{i=1}^{N}\frac{\nu(\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2 - \lambda}{((\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2+\lambda)\sqrt{\lambda}} \quad (13)$$

$$\frac{d\log p(\mathbf{z})}{d\sqrt{\nu}} = \Big[N\log\lambda - \sum_{i=1}^{N}\log((\mathbf{D}_i(\mathbf{z}-\mathbf{s}))^2+\lambda) + N\psi\left(\frac{1+\nu}{2}\right) - N\psi\frac{\nu}{2}\Big]\sqrt{\nu} \quad (14)$$

where $\psi(.)$ is the digamma function. The zeros of these functions give us the estimates of $\pm\sqrt{\lambda}$ and $\pm\sqrt{\nu}$. Although there is no closed-form solution, these can be obtained numerically using any unconstrained non-linear optimisation method[1], which requires the gradient vector of the objectives.

3.3 Recovery Algorithm

Our algorithm that implements the equations given in the previous section is given in Algorithm 1. Note that at each iteration of the algorithm, two smaller gradient descent problems have to be solved; namely one for λ, ν and one for $\mathbf{z}$. However, experiment suggests that it is not necessary to estimate the minimum with high accuracy. We notice that the inner loops do not require the entire convergence. It is sufficient to increase but not necessarily minimise the objective at each intermediate step.

Algorithm 1: Recovery algorithm.

1: Initialise the estimates $\mathbf{z}$
2: iterate until convergence: **do**
3: estimate σ^2 using (11)
4: iteratively update λ and ν in turn using definiton
5: (13) and (14), with the current estimate $\mathbf{z}$.
6: iterate to update $\mathbf{z}$ using (10)
7: **end**

4 EXPERIMENTS AND DISCUSSION

We design our experiments for both CS and SR-type W and we compare with the previous works in (A. Kabán and S. AliPitchay, 2011). We devise two hypotheses to investigate the role of the new prior and we test those using synthetic 1D and 2D signals and real MRI signals. Our hypotheses are the following:

- The quality of the recovered signal using the additional information is no worse than the one without the extra information provided that the extra information is *useful*. This is when the number of zero entries in the new form of the neighbourhood feature, i.e **Df** is larger than the number of zero entries in **Dz**, that is the generic feature that has not been given the extra similarity information.
- The fewer the edges in **f** (that is, the non-zeros in **Df**), the fewer measurements are sufficient for enabling a successful recovery.

Before we proceed with the experiments, we should mention the construction of the measurement matrix **W**. We study two different types: CS-type W is a random Gaussian matrix ($M \times N$) with iid entries. The SR-type W is a deterministic transformation that blurs and downsamples the image[2].

4.1 Illustrative 1D Experiments

In this section, we implement our recovery algorithm on the 1D data, derived from a spike signal[3] of size

[1] We made use of the efficient implementation available from http://www.kyb.tuebingen.mpg.de/bs/people/carl/code/minimize/

[2] Code to generate SR-type matrices can be found from http://www.robots.ox.ac.uk/∼elle/SRcode/ index.html

[3] Data taken from http://people.ee.duke.edu/∼lcarin/BCS.html

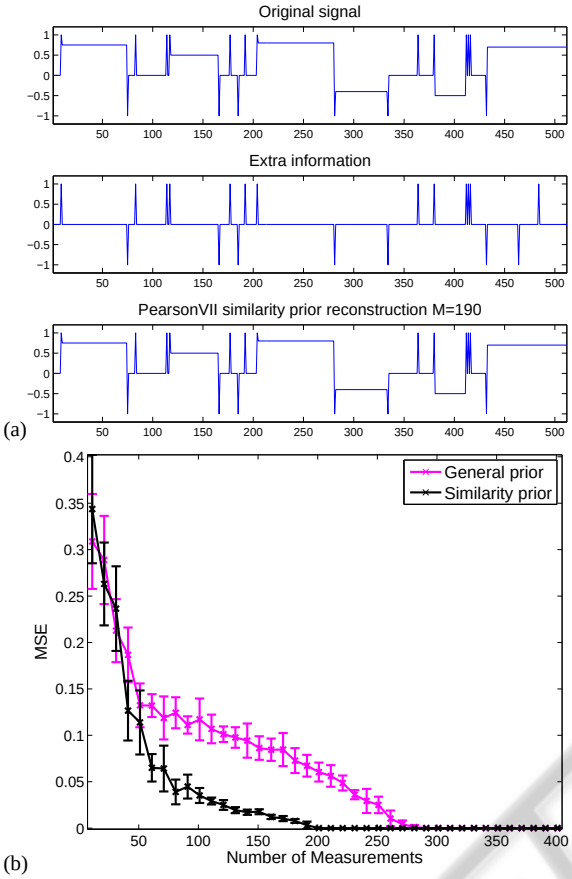


Figure 2: (a) The original spike signal; the extra similarity information; and an example of recovered signal from 190 measurements. (b) Comparing the MSE performance of 1D spike signal recovery with and without the extra information. The error bars are over 10 independent trials and the level of noise was σ=8e-5.

512x1 as shown in Figure 2(a). We proceed by plugging the extra signal into our image prior and varying the number of measurements using randomly generated measurement matrices **W** with i.i.d. Gaussian entries as in CS. The recovery results are summarised in Figure 2. We see our enhanced prior is capable to achieve a good recovery and has a lower mean square error (MSE) than the one without extra information.

We also examine the MSE performance as a function of the number of zero entries in the relevant feature vectors (i.e. **Df** in our case). Figure 3 shows MSE results when varying the number of zero entries by constructing variations on the signals. We see when the recovery algorithm received sufficient measurements, for example when M=250 in Figure 2, the role of the proposed *similarity prior* gradually reduces. In other words, this *similarity prior* is *useful* in massively underdetermined problems and provided that the given extra information has the characteristics described previously.

A widely used alternative way to set hyperparameters is cross-validation. It is therefore of interest how does the automated estimation of the hyper-parameters of our Pearson type VII based MRF compare to a cross-validation procedure. Next, we address this by looking at two aspects: MSE performance, and CPU time. We use the same spike signal for this purpose. For our comparison, we have chosen 5-folds cross validation method for estimating the hyper-parameters λ and ν and the noise variance is assumed to be known for this method. A sensible search range is pursued to avoid a long execution time as we are aware that this method can be extremely time-consuming if the search space is too large.

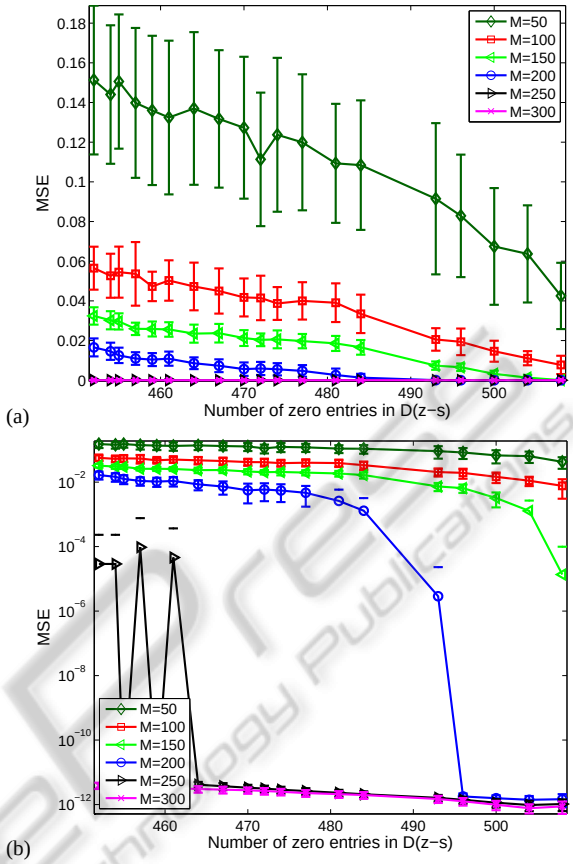


Figure 3: (a) Linear scale. (b) Log scale. MSE performance of 1D spike signal using the extra information. The number of zero entries in **D**(**z**-**s**) is varied. The error bars represent one standard error about the mean, from 50 independent trials. The level of noise was σ=8e-5.

Figure 4 shows the MSE performance and the associated values for the four levels of noise using the CS-type **W**. It is interesting to see that our fully automated parameter estimation turns out to be superior to 5-folds cross validation and it has fast convergence and much lower execution time.

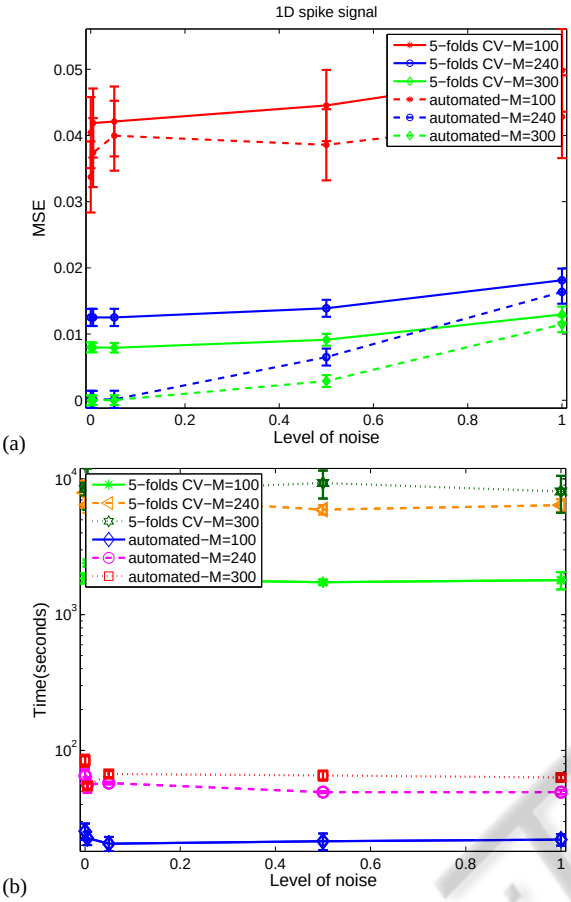


Figure 4: (a) Comparing the MSE performance of the fully automated Pearson type VII based MRF approach with the 5-folds cross validation, tested with four levels of noise (σ= 0.005, 0.05, 0.5, 1). (b) Cpu time performance against the same four levels of noise. We see that our automated estimation and recovery is significantly faster than the 5-folds cross validation method. The error bars are over 10 repeated trials for each level of noise. Three sets of measurements (M=100, 240, 300) have been tested for this accuracy comparison.

4.2 2D Experiments

Following the thorough understanding gained in the previous section about when the extra information is helpful on the spike signal test cases, we conducted experiments with both compressive sensing (CS) matrices where W contains random entries and also the classical super-resolution matrices where W consists of blur and down-sampling. In this set of experiments, we consider a motionless scene as the extra information. More precisely, the extra information that we employ in our similarity-prior consists of a change in the lighting of some area in the image.

We start by conducting the recovery algorithm on a synthetic data of size [50x50]. The noise variance σ tested in all experiments are set to a smaller range in order to tally the general noise in real data.

Figures 5 and 6 show examples of vastly underdetermined problems using the extra information for recovery in comparison with the previous prior devised in (A. Kabán and S. AliPitchay, 2011). The MSE performance results are given in Figure 7, and we see the MSE drops rapidly with increasing the measurement size. Figure 8 shows examples of recovered images from this process. We observe that the quality of the recovered image increases rapidly for all 5 levels of noise tested. This is in contrast with the recovery results from the general prior, which needs a lot more measurements to perform well.

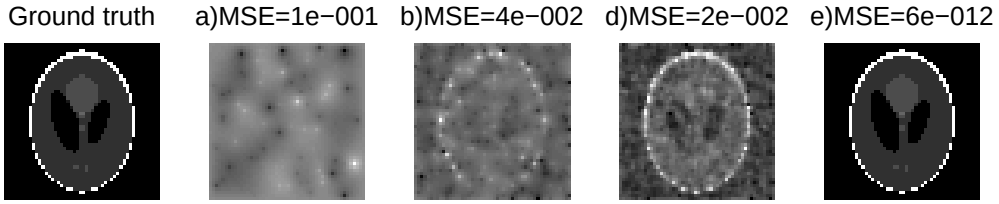
From these findings, the degree of similarity of the available extra information has a significant impact on the recovery from insufficient measurements. We find that without informative extra information the recovery algorithm does not perform well with such few measurements. The recovered signal and the MSE using the artificial *Phantom* data in figures 5 and 7 demonstrate that the fewer the edges in the difference image**f** the better the recovery, or the smaller the number of measurements needed for good recovery. This result validates our second hypothesis.

In the remainder of the experiments, we will now focus on image recovery using real image data of magnetic resonance imaging (MRI). We obtained this data from the Matlab database and we created the additional similarity information from it by changing the lighting of an area on the image.

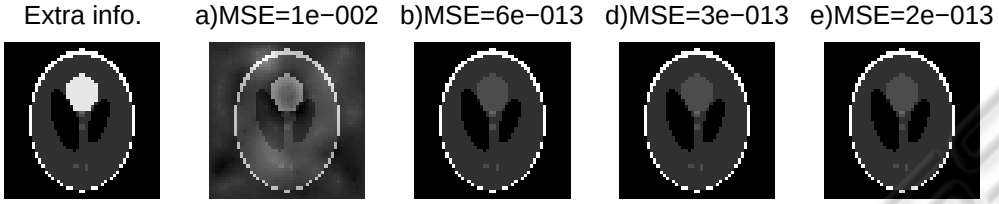
Next we validate our second hypothesis on a variety of MRI images and its lighting changes. The recovery results for both types of W are presented in figures 10 and 11. The MSE performance for the CS-type W is shown in figure 9. Interestingly, we observe that the log scale in that figure is in more direct correspondence with our visual perception rather than using the standard linear scale, and this will be seen by comparison to figures 10 and 11.

We observed that more than 6000 measurements are required for a good recovery without the extra information in this example. However, from these results we see that our similarity prior achieves high quality recovery from an order of magnitude less measurements. The recovered images are presented in figures 10 and 11 for visual comparison. Finally, we also show an example run of our automated parameter estimation algorithm in Figure 12 for completeness. As one would expect, the speed of convergence varies with the difficulty of the problem.

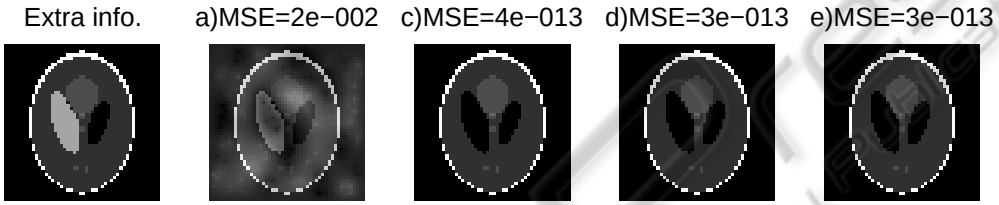
In closing, we should comment on the possibility of using other types of extra information for signal recovery. Throughout this paper we exploited the similarity created by a lighting change. Depending on the application domain, one might consider a small shift or rotation instead. However, we have seen that the key for the



Sample image recovery without extra information.

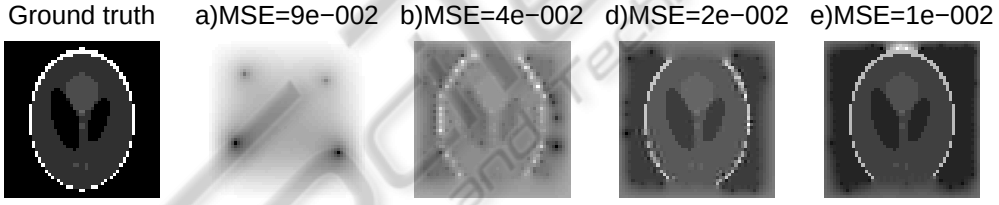


Sample image recovery using extra information of lighting change 1.



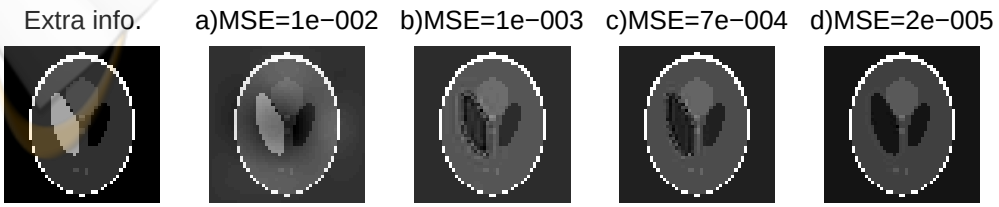
Sample image recovery using extra information of lighting change 2.

Figure 5: Example recovery of 2D synthetic data of size [50x50] in the case of using SR-type W, and given two slightly different light changes as extra similarity information. The number of measurements (M) are: a) M=60, b) 460, c) 510, d) 960, e) 1310. The additive noise level was σ=8e-5.



Sample image recovery without extra information.

Sample image recovery using extra information of lighting change 1.



Sample image recovery using extra information of lighting change 2.

Figure 6: Example recovery of 2D synthetic data of size [50x50] in the case of using SR-type W, and given two slightly different light changes as extra similarity information. The number of measurements (M) are: a) M=9, b) 441, c) 784, d) 1296, e) 1849. The additive noise level was σ=8e-7.

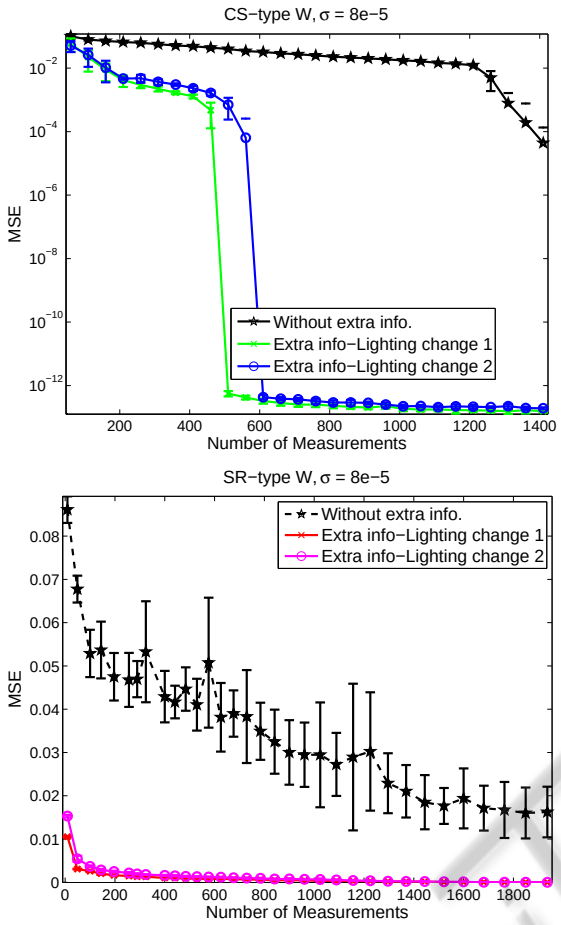


Figure 7: MSE performance of synthetic data [50x50] in comparison with the two types of extra information. Here, both types of W were tested and the noise standard deviation was σ=8e-5.

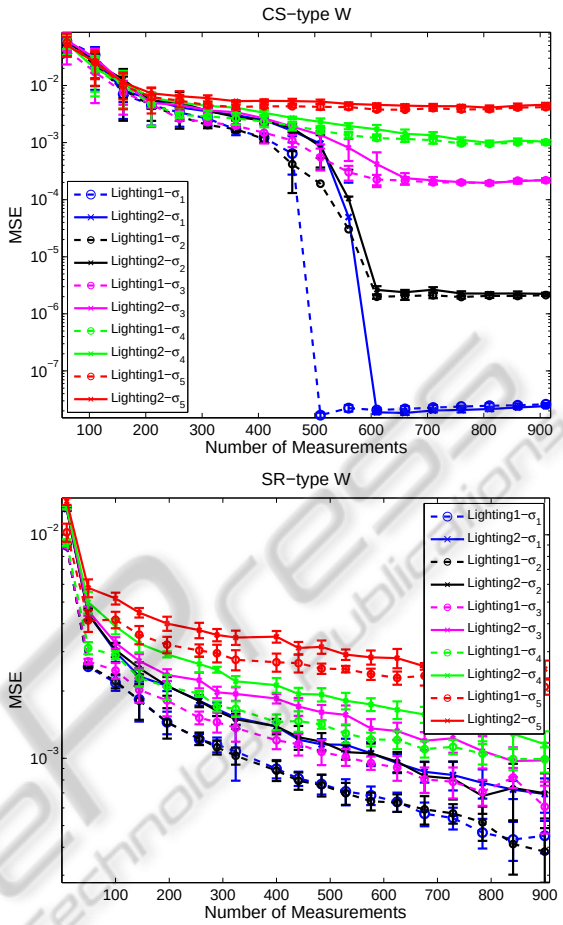


Figure 8: Recovery of a 50x50 size image from random measurements (top) and blurred and down-sampled measurement (bottom). The MSE is shown on log scale against varying the number of measurements, in 5 different levels of noise conditions. The noise levels were as follows. Top: $\sigma \in \{\sigma_1$=0.005, σ_2=0.05, σ_3=0.5, σ_4=1, σ_5=2$\}$; Bottom: $\{\sigma_1$=8e-5, σ_2=8e-4, σ_3=8e-3, σ_4=0.016, σ_4=0.032$\}$ — that is the previous noise levels were divided by 0.8 $\sqrt{N}$ to make the signal-to-noise ratios roughly the same for the two measurement matrix types.

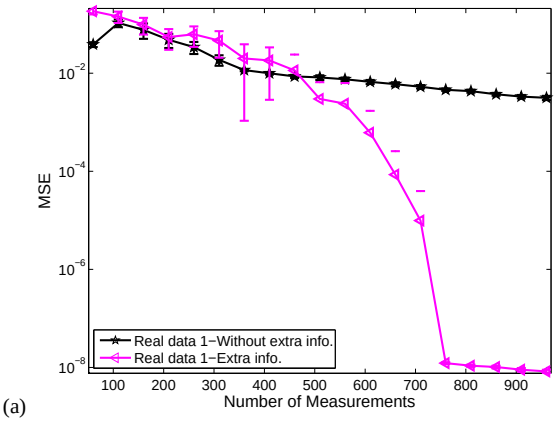
extra information to be useful in our similarity prior is that the difference image must have fewer edges than the original image. This is not the case with shifts or rotations. Therefore to make such extra information useful we would need to include an image registration model into the prior. This is subject to future work.

5 CONCLUSIONS

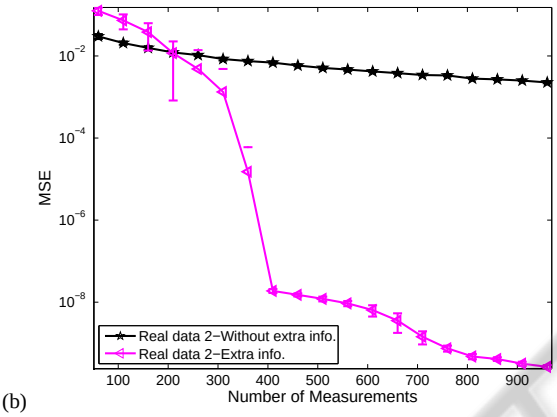
In this paper, we have formulated and employed a *similarity prior* based Pearson type VII Markov Random Field to include the similarity information between the scene of interest and a consecutive scene that has a lighting change. This prior enables us to recover the high resolution scene of interest from fewer measurements than a general-purpose prior would, and this can be applied, e.g. in medical imaging applications.

ACKNOWLEDGEMENTS

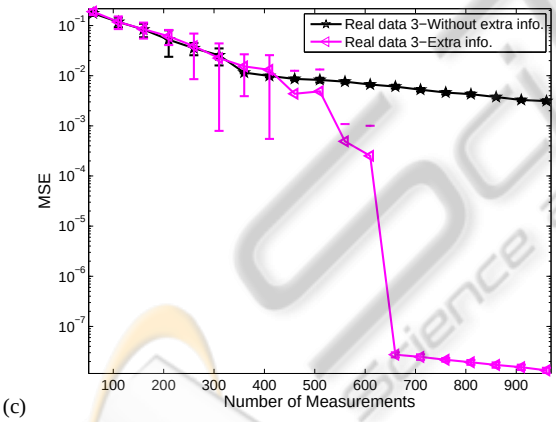
The first author wishes to thank *Universiti Sains Islam Malaysia* (USIM) and the Ministry of Higher Education of Malaysia (MOHE) for the support and facilities provided.



(a)

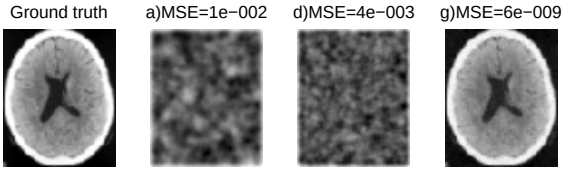


(b)

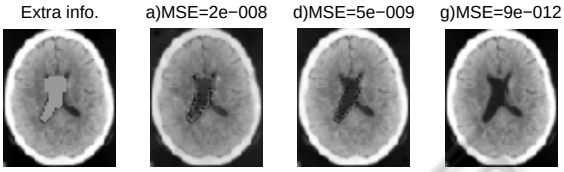


(c)

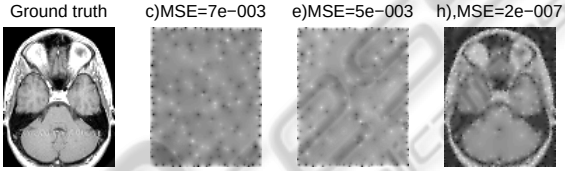
Figure 9: MSE performance of real MRI images of size (a)[70x57], (b) and (c) [100x80], in comparison with three types of extra information on the three different sets of data. CS-type W was used and the noise standard deviation was σ=8e-5.



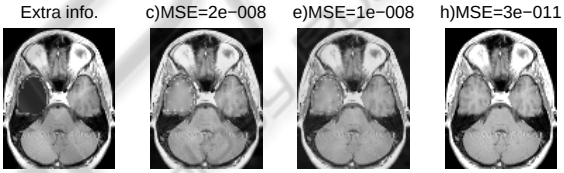
Sample image recovery of size [70x57] without extra information.



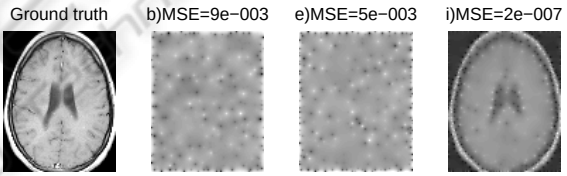
Sample image recovery of size [70x57] using extra information.



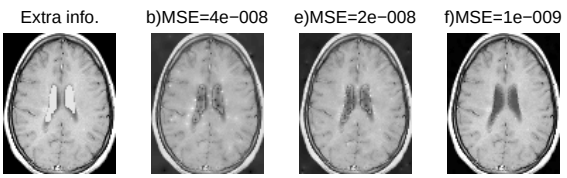
Sample image recovery of size [100x80] without extra information.



Sample image recovery of size [100x80] using extra information.



Sample image recovery of size [100x80] without extra information.



Sample image recovery of size [100x80] using extra information.

Figure 10: Examples of MRI image recovery in the case CS-type W, given a motionless consecutive frame with some contrast changes. The number of measurements (M) were: a) M=310, b) 460, c) 560, d) 610, e) 760, f) 1310, g) 3010, h) 5610 i) 7610 and additive noise with σ = 8e-5. The first two row refers to real data 1, the third row refers to real data 2 and the fifth row refers to real data 3.

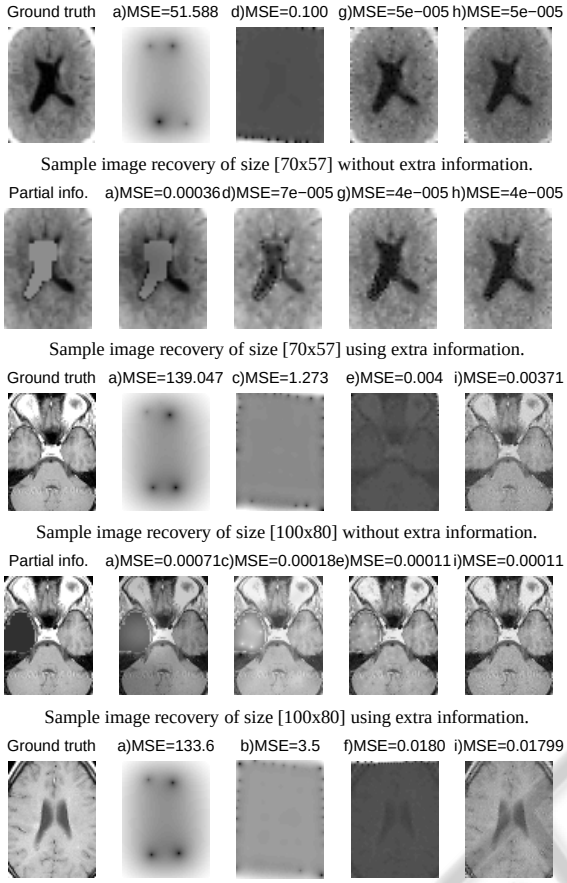


Figure 11: Examples of MRI image recovery in the case of SR-type W, given a motionless consecutive frame with some contrast changes. The number of measurements (M) were: a)M=6, b) 99, c) 154, d) 396, e) 918, f) 1462, g) 1505, h) 2000, i) 4234. The additive noise is σ=8e-5.

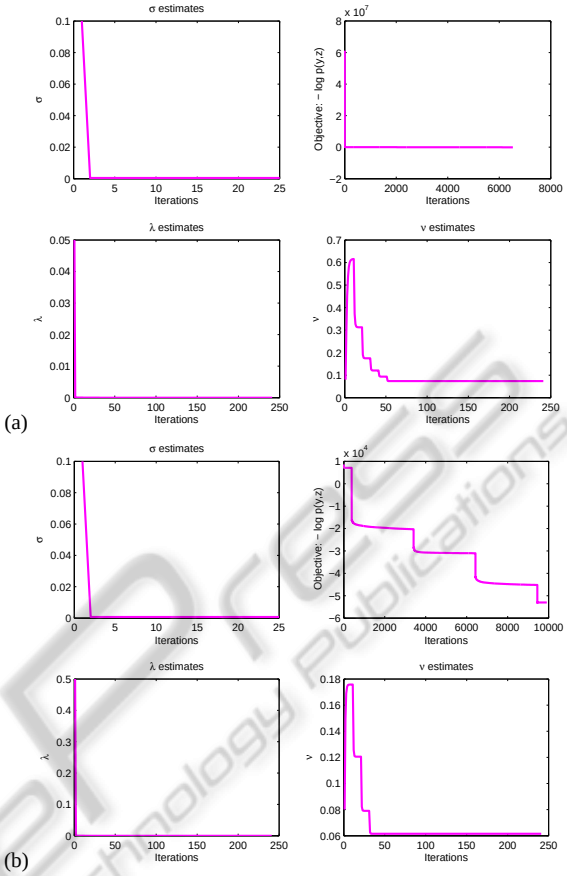


Figure 12: Example evolution of the hyper-parameter updates (σ, λ, ν) and objective function versus the number of iterations of the optimisation algorithm while recovering a 2D signal: (a) from M=460 random measurements; (b) from a blurred and down-sampled low resolution frame of M=144. In both experiments, the level of noise was σ=8e-5.

REFERENCES

A. Kabán and S. AliPitchay (2011). Single-frame image recovery using a pearson type vii mrf. In *Special issue of Neurocomputing on Machine Learning for Signal Processing*. accepted.

D. L. Donoho (2006). Compressed sensing. *IEEE Trans. Information Theory*, 52(4):1289–1306.

E. Candes, J. Romberg, and T. Tao (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Information Theory*, 52(2):489 509.

H. He and L. P. Kondi (2003). Map based resolution enhancement of video sequences using a huber-markov random field image prior model. In *IEEE Conference of Image Processing*, pages 933–936.

H. He and L. P. Kondi (2004). Choice of threshold of the huber-markov prior in map based video resolution enhancement. In *IEEE Electrical and Computer Engineering Canadian Conference*, volume 2, pages 801–804.

JCR. Giraldo, JD. Trzasko, S. Leng, CH. McCollough, and A. Manduca (2010). Non-convex prior image constrained compressed sensing (nc-piccs). In *Proc. of SPIE : Physics of Medical Imaging*, volume 7622.

L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zissermann (2007). Bayesian methods for image super-resolution. *The Computer Journal*.

N. Vaswani and W. Lu (2010). Modified-cs: Modifying compressive sensing for problems with partially known support. *IEEE Trans. on Signal Processing*, 58(9).

RC. Hardie and KJ. Barnard (1997). Joint map registration and high-resolution image estimation using a se-

quence of undersampled images. *IEEE Trans. Image Processing*, 6(12):621–633.

R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde (2010). Model-based compressive sensing sensing. *IEEE Trans. Information Theory*, 56:1982–2001.

S. Ji, Y. Xue, and L. Carin (2008). Bayesian compressive sensing. *IEEE Trans. Signal Processing*, 56(6):2346–2356.

W. Lu and N. Vaswani (2011). Regularized modified bpdn for noisy sparse reconstruction with partial erroneous support and signal value knowledge. *IEEE Trans. on Signal Processing*, To Appear.