**Proceedings Paper:**

Hotrakool, W. and Abhayaratne, C. orcid.org/0000-0002-2799-7395 (2015) An optimal learning parameter for running Gaussian-based referenced compressive sensing. In: 2nd Proceedings of IET International Conference on Intelligent Signal Processing 2015 (ISP). 2nd IET International Conference on Intelligent Signal Processing 2015 (ISP), 01-02 Dec 2015, London, UK. IET . ISBN 9781785611360

https://doi.org/10.1049/cp.2015.1759

# An Optimal Learning Parameter for Running Gaussian-based Referenced Compressive Sensing

**Wattanit Hotrakool[1], Charith Abhayaratne[2]**

Department of Electronic and Electrical Engineering, The University of Sheffield, United Kingdom
[1] w.hotrakool@sheffield.ac.uk, [2]c.ahayaratne@sheffield.ac.uk

**Keywords:** compressive sensing, reconstruction, side-information

## Abstract

One of the approaches to exploit temporal redundancy in compressive sensing reconstruction of spatio-temporal signals is the Running Gaussian-based Referenced Compressive Sensing. It uses the weighted-average of all prior reconstructed instances as a reference to reconstruct the next instance with high accuracy. The performance of this approach depends on the weight called learning parameter. This work studies the relationship between the learning parameter and the reconstruction accuracy. We show that the small value of the learning parameter is more suitable for natural signals with dynamic sparse supports. We also propose a dynamic optimal learning parameter that provides good reconstruction accuracy for all signals. Out experimental results show that the proposed optimal learning parameter outperforms all fixed values of learning parameter in natural video sequences reconstruction.

## 1 Introduction

Compressive sensing (CS) is an acquisition framework which enables the reconstruction of a full-length signal from its under-sampled measurements. Compressive sensing, introduced in [1, 2], has become of interest in various fields of research because of its ability to perform sub-Nyquist sampling.

There are two major components for a successful compressive sensing. The first component is the sensing operation. Given an $n$-dimensional signal $\mathbf{x} \in \mathbb{R}^n$, compressive measurements $\mathbf{y} \in \mathbb{R}^m$, $m \ll n$, is obtained using a sensing operator:

$$\mathbf{y} = \mathbf{\Phi\Psi x}, \tag{1}$$

where $\mathbf{\Psi}$ is a sparse basis for $\mathbf{x}$ and $\mathbf{\Phi}$ is an under-sample operator. Together, they form a sensing matrix $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{A} = \mathbf{\Phi\Psi}$.

The second component of compressive sensing is the reconstruction operation that reconstructs a full-length signal $\hat{\mathbf{x}} \in \mathbb{R}^n$ from the compressive measurements $\mathbf{y}$. The linear system in Eq. (1) is undersampled, to solve for $\hat{\mathbf{x}}$, the reconstruction is done using a convex optimisation. If the signal $\mathbf{\Psi x}$ is sparse, the best approximation $\hat{\mathbf{x}}$ is the sparsest solution in the feasible set of solution. Given that the sensing matrix $\mathbf{A}$ satisfy the incoherence property [3] and that $\mathbf{\Psi x}$ is sufficiently sparse, we can obtain $\hat{\mathbf{x}} \approx \mathbf{x}$ with high probability. The optimal solution to the optimisation problem

$$\min_{\hat{\mathbf{x}}} \|\mathbf{\Psi}\hat{\mathbf{x}}\|_1 \text{ subject to } \mathbf{A}\hat{\mathbf{x}} = \mathbf{y} \tag{2}$$

is the sparsest solution $\hat{\mathbf{x}}$.

However, in practice, reconstruction of signals by purely maximising their sparsity has limitations. Most real-world signals are not sufficiently sparse as they contain many small non-zero elements. Maximising of sparsity alone does not work well with these small non-zero elements and tends to result in reconstruction error. To improve the reconstruction accuracy, *a priori* information has been exploited in reconstruction operator. To date, several reconstruction methods using side-information have been proposed. They exploit the fact that many characteristics of a signal are shared between its neighbours. This is true in most natural signals, particularly in video sequences [4], multi-view imaging [5], and clinical imaging such as the magnetic resonance imaging (MRI) [6, 7]. The use of side-information generally improve the accuracy of the reconstruction results as well as reducing the number of measurements required. Several methods incorporate side-information successfully including sparse support estimation [8], model-based CS [9], Kalman-filtered CS[10] and group reconstruction[5]. The video sequences, in particular, are of special interest, because they contain high level of temporal redundancy that can be used easily as side-information. This redundancy is successfully exploited in Distributed Compressive Video Sensing (DCVS) [11].

In our previous works [12, 13], we have proposed the generalised approach to exploit the temporal redundancy. This method, which we refer to as Referenced Compressive Sensing, tries to minimise the error between a signal of interest to a reference, an arbitrary signal that is known to be close to the signal of interest. Also, we have shown that the reconstruction error is limited by the distance from the reference to the signal. We also proposed a running Gaussian-based reference estimation that improves the performance of the Referenced CS. This estimator works by using the weighted average of all reconstructed frames as a reference. The weight, which is referred to as the learning parameter $0 \leq \alpha \leq 1$, governs to characteristic of the estimated reference.

To date, however, there is no study on the relationship between the learning rate $\alpha$ and the reconstruction accuracy. The main contribution of this work is to study such relationship, as well as establish the optimal learning parameter. Our main result finds that the learning parameter impacts the reconstruction differently depending on the changes of the locations of sparse supports (the positions of non-zero coefficients). Also the dynamic optimal learning parameter, which provides an optimal performance regardless of supports' changes, is shown to work best in natural video sequence reconstruction.

## 2 Running Gaussian-based Referenced Compressive Sensing – Revisited

As introduced in our previous works [12], the accuracy of compressive sensing reconstruction can be improved greatly by exploiting the redundancy between signals. This is done by minimising the error between a signal and its correlated reference, which is an arbitrary signal very close to the signal of interest. We refer to this reconstruction approach as Referenced Compressive Sensing (Referenced CS).

Consider a large signal, such as an image or a video sequence, that can be viewed as a collection, denoted $C$, of several smaller signals. There are $k$ signals $\mathbf{x}_i \in C, \mathbf{x}_i \in \mathbb{R}^n, i = 1, 2, ..., k$, where $n$ is the length of each signal $\mathbf{x}_i$. Here we define the correlated reference $\mathbf{r}$.

**Definition 1.** *For any signal* $\mathbf{x}$*, a correlated reference* $\mathbf{r}$ *of* $\mathbf{x}$ *is a signal such that* $\mathbf{r} \in \mathbb{R}^n$ *and*

$$\|\mathbf{r} - \mathbf{x}\|_2 \leq \epsilon, \tag{3}$$

*for a small* $0 < \epsilon \ll \|\mathbf{x}\|$.

The distance between the reference $\mathbf{r}$ and $\mathbf{x}$ is denoted $\delta = \|\mathbf{r} - \mathbf{x}\|_2$. The pair of $\mathbf{x}$ and $\mathbf{r}$ can be anything, *e.g.*, images of the same scene, different rows of the same image, for examples. In this work, however, our focus is on the spatio-temporal signals such as video sequences. In this type of signal, the pair of $\mathbf{x}$ and $\mathbf{r}$ can be different instances of the same sequence.

In [12], we shows that the reconstructed signal $\hat{\mathbf{x}}$ of $\mathbf{x}$ has a guaranteed bound described by the reference distance $\delta$.

**Proposition 1.** *Given a sensing operator* $\mathbf{A} \in \mathbb{R}^{m \times n}$*, a compressed measurement* $\mathbf{y} \in \mathbb{R}^m, \mathbf{y} = \mathbf{A}\mathbf{x}$*, and a correlated reference* $\mathbf{r}$*, the least* $l_1$*-norm reconstruction* $\hat{\mathbf{x}}_1$*, which is the solution of*

$$\min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}} - \mathbf{r}\|_1 \text{ subject to } \mathbf{A}\hat{\mathbf{x}} = \mathbf{y}, \tag{4}$$

*satisfies*

$$\|\hat{\mathbf{x}}_1 - \mathbf{x}\|_2 \leq 2\delta. \tag{5}$$

The proof of the *Proposition 1* can be found in [12].

Donoho's Lemma 3.1 in [1] holds that for any $\hat{\mathbf{x}} \in \hat{\mathbf{X}}_{\mathbf{A},\mathbf{y}}$,

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq 2E_n(\hat{\mathbf{X}}_{\mathbf{A},\mathbf{y}}), \tag{6}$$

where $E_n(\hat{\mathbf{X}}_{\mathbf{A},\mathbf{y}})$ denotes the optimal solution in the feasible set $\hat{\mathbf{X}}_{\mathbf{A},\mathbf{y}}$. This implies that the Referenced CS solution in *Proposition 1* is at worst equivalent to the optimal solution of $l_1$-minimisation. The best performance, however, depends on the reference distance $\delta$. Therefore it is essential to choose the reference in that fashion that minimise $\delta$.

One easy way to exploit the redundancy between frames in a video sequence is to use a reconstructed frame as the reference for reconstructing the next frame, *i.e.*, set $\mathbf{r}_t = \hat{\mathbf{x}}_{t-1}$ at instance $t$. The problem with this approach is that by doing so, the reconstruction error are propagated from frame to frame. This results in the accumulated amount of error in the reconstructed sequence. In [13], we proposed the Running Gaussian-based reference estimator to cope with this error propagation issue. This estimator, inspired by Running Gaussian Average and Gaussian Mixture techniques, uses the combination of all the reconstructed frames of the same sequence to estimate the references. Running Gaussian estimator models the reference $\mathbf{r}$ as a vector of random variables drawn from normal distribution, *i.e.*,

$$\mathbf{r} = \{r_j : r_j = \mathcal{N}(\mu, \sigma^2)\}, \tag{7}$$

where $\mu$ and $\sigma^2$ are the mean and the variance of the distribution. We defined the update rule of $\mathbf{r}_t$, the reference at instance $t$, as

$$\mathbf{r}_t = \alpha\hat{\mathbf{x}}_{t-1} + (1 - \alpha)\mathbf{r}_{t-1}. \tag{8}$$

The parameter $\alpha$ is called a learning parameter. It governs the update rate of the estimation. The $\alpha \to 1$ gives the system that is more sensitive to the changes of signal's contents, making the reference more resemble to the latest reconstruction. The low $\alpha \to 0$, however, makes the system more robust to reconstruction error propagation. However, currently there is no study on the relationship between the learning parameter and the reconstruction performance. The learning parameter, so far, is fixed empirically to a scalar, and there is no optimal learning parameter for any arbitrary signal.

## 3 Learning Parameter for Referenced Compressive Sensing

This section presents our main results. Here, we study the relationship between the learning parameter and the reconstruction accuracy. Also, the optimal learning parameter will be defined such that it negates the propagated reconstruction error from the estimated reference.

To study the relationship between the learning parameter and the accuracy, we employ Monte Carlo method. Here we compute the reconstruction error in term of Peak Signal-to-Noise Ratio (PSNR) from the sets of constructed sparse sequences $\mathbf{X}$. Each sequence $\mathbf{x} \in \mathbf{X}$ is a sparse signal with $k$ non-zero elements,*i.e.*, $k$-sparse, for a small $k$. Both the locations and magnitudes of the sparse supports of $\mathbf{x}_1$, the first instance of $\mathbf{X}$, are drawn from uniformly random process. To generate other instances $\mathbf{x}_i \in \mathbf{X}, i > 1$, while maintaining their likeliness with $\mathbf{x}_1$, we employ the following procedure:

1. A new support is randomly added to the supports set of $\mathbf{x}_i$. Its magnitude is drawn randomly.
2. An existing support of $\mathbf{x}_i$ is randomly removed. Its magnitude is reset to zero.
3. Each element of $\mathbf{x}_i$ is multiplied by a random gain $0.9 \leq \gamma \leq 1.1$.
4. Each support, along with its magnitude, of $\mathbf{x}_i$ is randomly shifted.

Each sparse set $\mathbf{X}$ is then compressively sampled and reconstructed using Referenced CS in Eq. (4). The reference used for the reconstruction is estimated using the Running Gaussian estimator in Eq. (8) using varying value of the learning parameter $\alpha$. Two groups of $\mathbf{X}$, each contains the total of 1000 sets of $\mathbf{X}$, are employed in Monte Carlo. The first group of $\mathbf{X}$ is created without using the last procedure, *i.e.*, no shift in the locations of the sparse supports. Figure 1 shows the scatter plot between the PSNR and the learning parameter $\alpha$ of this group. It can be seen using a regression line that, when the sparse supports are stationary, the large value of $\alpha$ provides the results with highest PSNR with highest probability. The use of the naive reference ($\alpha = 1$) also provides a very good accuracy, thus the use of Running Gaussian estimator is trivial.

The second group of $\mathbf{X}$ is created with the random support shift procedure. In can be seen in Figure 2 that the situation is reverse when the sparse supports are no longer stationary. In this case, the use of small values of $\alpha$ gives better reconstruction accuracy than the large values. The middle range of $\alpha$ provides a middle ground for both signals with stationary and non-stationary supports.

Instead of using a fixed learning parameter $\alpha$, it is possible to use a dynamic rate. Given a collection of spatio-temporal signal $C$, we can express the reconstructed signal $\hat{\mathbf{x}}_t$ of $\mathbf{x}_t \in C$ as

$$\hat{\mathbf{x}}_t = \mathbf{x}_t + \mathbf{e}_t, \tag{9}$$

where $\mathbf{e}_t$ is the reconstruction error. We assume that $\mathbf{e}_t$ is a vector of random variable drawn from a random process $\mathbf{E}$ of some unspecified distribution. Suppose we require the reference $\mathbf{r}_{t+1}$ to be an average of the first $t$ instances of $\hat{\mathbf{x}}$, *i.e.*,

$$\mathbf{r}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} \hat{\mathbf{x}}_i. \tag{10}$$

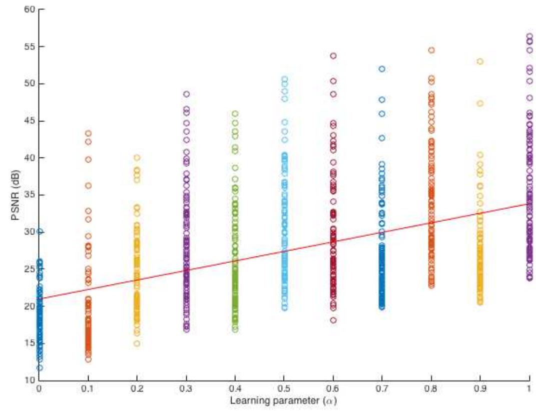


Figure 1: Relationship between the learning rate $\alpha$ and reconstruction accuracy when sparse supports are stationary

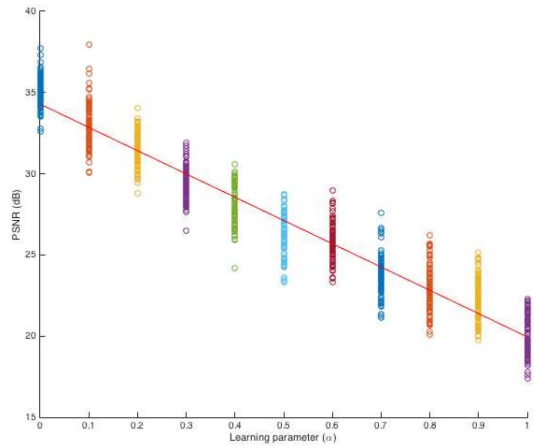


Figure 2: Relationship between the learning rate $\alpha$ and reconstruction accuracy when sparse supports are dynamic

Since

$$\mathbf{r}_t = \frac{1}{t-1}\sum_{i=1}^{t-1} \hat{\mathbf{x}}_i, \tag{11}$$

we can derive that

$$\mathbf{r}_{t+1} = \frac{1}{t}\hat{\mathbf{x}}_t + \frac{t-1}{t}\mathbf{r}_t. \tag{12}$$

Thus, by setting $\alpha = \frac{1}{t}$ for any value of $t$ makes $\mathbf{r}_t$ to be the average of the first $t-1$ instances of $\hat{\mathbf{x}}$.

The learning parameter $\alpha = 1/t$ is optimal. By expressing Eq. (11) in terms of Eq. (9), we can see that

$$\mathbf{r}_t = \frac{1}{t-1}[(\mathbf{x}_1 + \mathbf{e}_1) + (\mathbf{x}_2 + \mathbf{e}_2) + \cdots + (\mathbf{x}_{t-1} + \mathbf{e}_{t-1})] \tag{13}$$

$$= \frac{1}{t-1}[(\mathbf{x}_1 + \cdots + \mathbf{x}_{t-1}) + (\mathbf{e}_1 + \cdots + \mathbf{e}_{t-1})]. \tag{14}$$

Since each $\mathbf{e}$ term is drawn from a random process, by Central Limit Theorem, when $t$ is large, the distribution of the summation of $\mathbf{e}$ becomes a normal distribution. Thus, we obtain

$$\mathbf{r}_t \equiv \frac{\sum_{t=1}^{t-1} \mathbf{x}_i}{t-1} + \frac{(t-1)\mathcal{N}(0,\sigma^2)}{t-1} \tag{15}$$

$$\equiv \bar{\mathbf{x}}_{t-1} + \mathcal{N}(0,\sigma^2), \tag{16}$$

where $\bar{\mathbf{x}}_{t-1}$ is the average of the first $t-1$ instances. By using Monte Carlo, the average PSNR when using this optimal learning parameter is 28.23 dB for the stationary supports group and 29.72 dB for the dynamic group.

## 4 Experimental Results

In this section, the proposed reconstruction method is employed to reconstruct real video sequences. The dataset used in this experiment consists of 14 video sequences. These sequences are grouped into 3 categories: low activity sequences, medium activity sequences, and high activity sequences. Low activity sequences contains very low amount of motion and are virtually static. In high activity sequences, on the other hand, the amount of motion is significant and is the dominant feature of the sequences. The medium activity sequences have the natural amount of motion between these two extremes.

The sequences are compressively sampled and reconstructed. The reconstruction is done by solving the optimisation problem in Eq. (4). Several strategies for choosing the correlated reference $\mathbf{r}$ are compared between each other in this experiment. The most simplest reference is the naive reference that uses the immediate reconstructed frame as a reference for the next frame, *i.e.*, $\mathbf{r}_t^{naive} = \hat{\mathbf{x}}_{t-1}$, at time $t$. The Running Gaussian references are estimated using fixed learning parameter $\alpha = 0.1$, $\alpha = 0.3$, and $\alpha = 0.5$, as well as the proposed adaptive optimal learning parameter $\alpha = 1/t$ at time $t$. The reconstructed results using these references are also compared with the reconstruction using lossless references. The lossless reference is the controlled benchmark, obtained directly from the full-length data without compressive sensing, *i.e.*, $\mathbf{r}_t^{lossless} = \mathbf{x}_{t-1}$. Such reference is of course unavailable in practice and is shown here only for comparison purpose.

Figure 3 shows the examples of the reconstructed sequences. Each row in Figure 3 shows the results obtained using a different kind of references. The sequences in the first and second columns are the examples of low activity sequence. The third and forth columns are the examples of medium activity sequences, while the last column shows the examples of high activity sequences. It is clear that, in all sequences, the reconstructions using references with optimal learning parameter have much better visual quality than the results using other references. It also shows that the results obtained from $\alpha = 0.1$ references are better than those obtained from $\alpha = 0.5$ references, and that the results using naive references demonstrate the lowest visual quality. This follows the discussion in Section 3 that when the locations of signal's supports are not stationary, the small value of $\alpha$ provides the most robust reconstruction results. Most supports of natural sequences are not stationary, particularly the supports of high frequency components, thus the small $\alpha$ provides the results with more accuracy than the larger ones and the naive reference (which is $\alpha = 1$). This observation is verified by Table 1, which shows the peak signal-to-noise ratio (PSNR) of each reconstructed sequence using each reference. This table confirms that the optimal parameter outperforms all fixed learning parameters in natural sequences reconstruction. It also confirms that the small $\alpha$ outperforms larger values of $\alpha$. Also, it shows that these effects of the learning parameter are more prominent when the activity level in the sequence is higher. As such, the difference in reconstruction quality using different learning parameter for references can be observe more easily in high activity sequences than in low activity sequences.

Table 1: Peak Signal-to-Noise Ratio of reconstructed video sequences using various types of references

| | Lossless | Naive | $\alpha = 0.5$ | $\alpha = 0.3$ | $\alpha = 0.1$ | optimal |
|---|---|---|---|---|---|---|
| Low activity sequences | | | | | | |
| 1 | 46.16 | 34.79 | 34.96 | 35.20 | 35.43 | **35.59** |
| 2 | 47.76 | 35.17 | 35.48 | 36.69 | 36.94 | **36.94** |
| 3 | 44.64 | 25.75 | 25.78 | 25.78 | 25.89 | **32.87** |
| 4 | 40.13 | 34.06 | 34.58 | 35.19 | 35.93 | **37.16** |
| 5 | 46.05 | 37.18 | 37.42 | 38.05 | 38.87 | **39.24** |
| Medium activity sequences | | | | | | |
| 6 | 37.53 | 31.53 | 31.36 | 32.37 | 32.50 | **33.56** |
| 7 | 35.37 | 29.66 | 29.96 | 30.62 | 31.24 | **32.11** |
| 8 | 41.08 | 32.20 | 32.70 | 32.81 | 33.08 | **33.56** |
| 9 | 44.78 | 34.63 | 34.73 | 35.75 | 35.79 | **36.09** |
| 10 | 42.89 | 32.31 | 32.45 | 32.55 | 33.06 | **33.11** |
| 11 | 41.98 | 31.55 | 31.84 | 32.57 | 32.74 | **33.17** |
| High activity sequences | | | | | | |
| 12 | 35.88 | 29.54 | 29.73 | 30.55 | 31.05 | **32.04** |
| 13 | 36.56 | 29.60 | 29.76 | 30.36 | 30.82 | **32.61** |
| 14 | 41.30 | 31.85 | 32.06 | 32.76 | 32.83 | **33.18** |

## 5 Conclusions

In this paper, we have discussed the relationship between the reconstruction accuracy and the learning parameter in the running Gaussian-based Referenced Compressive Sensing. We have shown that the effect of the learning parameter depends on the changes of the locations of sparse supports. That is, the large value of learning parameter is suitable for the signals with stationary supports, whereas the smaller values work better with the signals with dynamic supports. As most natural signals have dynamic supports, the small learning parameters

Low activity | Low activity | Medium activity | Medium activity | high activity

(a) Lossless reference

(b) Optimal reference

(c) Reference with $\alpha = 0.1$

(d) Reference with $\alpha = 0.3$

(e) Reference with $\alpha = 0.5$

(f) Naive reference

Figure 3: Examples of reconstructed sequences using various types of references

work better with such natural signals. We also defined the optimal learning parameter with the aim to eliminate the propagation of reconstruction error in the reference. This optimal learning parameter is shown to outperform any fixed values of parameter in natural video sequences reconstruction.

## References


[1] D.L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[2] 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, no. 2, pp. 489–509, February 2006.

[3] E.J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9–10, pp. 589–592, 2008.

[4] G. Warnell, D. Reddy, and R. Chellappa, "Adaptive rate compressive sensing for background subtraction," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2012, pp. 1477–1480, IEEE.

[5] V. Thirumalai and P. Frossard, "Joint reconstruction of multi-view compressed images," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1969–1981, 2012.

[6] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72 –82, March 2008.

[7] W. Hotrakool and C. Abhayaratne, "Efficient reconstruction of functional magnetic resonance imaging (fMRI) data from compressive measurements based on correlated reference," in *10th IMA International Conference on Mathematics in Signal Processing*, December 2014.

[8] J. Scarlett, J. S. Evans, and S. Dey, "Compressed Sensing With Prior Information: Information-Theoretic Limits and Practical Decoders," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 427–439, Jan. 2013.

[9] R.G. Baraniuk, V. Cevher, M.F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[10] N. Vaswani, "Kalman filtered compressed sensing," in *15th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2008, pp. 893–896.

[11] Li-Wei Kang and Chun-Shien Lu, "Distributed compressive video sensing," *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009.

[12] W. Hotrakool and C. Abhayaratne, "Fast compressed sensing reconstruction using the least squares and signal correlation," in *1st IET Intelligent Signal Processing Conference (ISP)*, December 2013, pp. 1.1–1.1(1).

[13] W. Hotrakool and C. Abhayaratne, "Running Gaussian Average-based reconstruction for video compressed sensing," in *2014 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2014.