Physics-Constrained Dictionary Learning for Selective Laser Melting Process Monitoring

Yanglong Lu, Yan Wang
Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

Abstract

Compressed sensing (CS) as a new data acquisition technique has been applied to monitor manufacturing processes. With a few measurements, sparse coefficient vectors can be recovered by solving an inverse problem and original signals can be reconstructed. Dictionary learning methods have been developed and applied in combination with CS to improve the sparsity level of the recovered coefficient vectors. In this work, a physics-constrained dictionary learning approach is proposed to solve both of reconstruction and classification problems by optimizing measurement, basis, and classification matrices simultaneously with the considerations of the application-specific restrictions. It is applied in image acquisitions in selective laser melting (SLM). The proposed approach includes the optimization in two steps. In the first stage, with the basis matrix fixed, the measurement matrix is optimized by determining the pixel locations for sampling in each image. The optimized measurement matrix only includes one non-zero entry in each row. The optimization of pixel locations is solved based on a constrained FrameSense algorithm. In the second stage, with the measurement matrix fixed, the basis and classification matrices are optimized based on the K-SVD algorithm. With the optimized basis matrix, the coefficient vector can be recovered with CS. The original signal can be reconstructed by the linear combination of the basis matrix and the recovered coefficient vector. The original signal can also be classified to identify different machine states by the linear combination of the classification matrix and the coefficient vector.

Keywords
Compressed sensing, Dictionary learning, Sparse coding, Data compression, Manufacturing process monitoring

1. Introduction
Sensors play an important role in modern manufacturing process monitoring to ensure product qualities. However, in-situ monitoring of complex manufacturing systems can be costly. Because of the bandwidth limitation of communication for the volume of transmitted data, there is a practical need to develop a scheme to reduce the amount of data collection without sacrificing the information exchanged. In the most recent decade, compressed sensing (CS) and dictionary learning approaches have been developed to improve the efficiency of data collection. CS [1, 2] allows us to reduce the amount of data collection by taking advantage of the sparsity of coefficient vector in reciprocal space. When the original signal is represented in a discrete form as vector $\boldsymbol{s} \in \mathbb{R}^N$. It can be represented in the reciprocal space via transformations as $\boldsymbol{s} = \boldsymbol{\Psi\gamma}$, where $\boldsymbol{\Psi} \in \mathbb{R}^{N\times N}$ is the transformation or basis matrix and $\boldsymbol{\gamma} \in \mathbb{R}^N$ is the vector of coefficients. When the signal is projected into the M-dimensional measurement subspace (M<N) with measurement matrix $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ as $\boldsymbol{y} = \boldsymbol{\Phi s}$, the original signal $\boldsymbol{s}$ can be recovered from the measured data $\boldsymbol{y}$ by solving the inverse problem based on the linear equations $\boldsymbol{y} = \boldsymbol{\Phi s} = \boldsymbol{\Phi\Psi\gamma}$. The basis matrix $\boldsymbol{\Psi}$ is usually predefined from some known transformation such as discrete cosine transformation, wavelet transformation, or some random matrices that satisfy the restricted isometry property. The combination of measurement and basis matrices $\boldsymbol{\Phi\Psi}$ is also referred to as dictionary. Dictionary learning methods have been developed to further improve the sparsity of coefficient vectors and the compression ratio of CS by customizing the basis matrix according to specific signal types. The learning process is formulated as an optimization problem where the entries of measurement and basis matrices are optimized to minimize the CS recovery error.

To further improve the efficiency of data collection in physical applications, a physics-constrained dictionary learning method was recently proposed and it has been demonstrated to compress roller bearing vibration signals [3]. The measurement and basis matrices were optimized simultaneously. There is only one non-zero entry in each row of the measurement matrix and the index of the non-zero entry indicates the time stamp of the stored data point. The physical constraint to indicate the maximum sampling rate was also incorporated to minimize the redundant information collection. With a few stored data points, the original vibration signals can be reconstructed. In this paper, the physics-constrained dictionary learning method is extended to solve the classification problems of machine health diagnostics. The measurement, basis and classification matrices are optimized simultaneously. With a few collected data points, the sparse coefficient vector is recovered with CS. The coefficient vector is then used to reconstruct the complete original signal by multiplying the basis matrix and obtain the class label by multiplying the classification matrix respectively. This new process monitoring and diagnostics method is demonstrated with selective laser melting (SLM) process, where low-resolution optical images of melt pool are collected to reconstruct high-resolution images and identify different machine states.

In the remainder of the paper, the background of dictionary learning methods and their applications is given in Section 2. The proposed physics-constrained dictionary learning algorithm is described in Section 3. The demonstration for the SLM process monitoring is given in Section 4.

2. Background

Various dictionary learning methods [4] have been developed to search for the sparsest representation of signals. The purpose is to find the optimal dictionary so that the sparsity is maximized for a specific type of signals. As a result, the original signals can be represented in a form of linear combinations of the learned dictionary and the sparse vector of coefficients. Some commonly used dictionary learning algorithms include the method of optimal directions (MOD) [5], K-SVD [6], the online dictionary learning [7] and others. Approaches to design the measurement matrix and the basis matrix individually were also developed. Duarte-Carvajalino and Sapiro [8] simultaneously optimized the measurement matrix and basis matrix with a new scheme called coupled-KSVD. The incoherence between the measurement matrix and basis matrix is improved which results in the better reconstruction performance. Bai et al. [9] further improved the framework with analytical solutions to update the measurement and basis matrices. It was shown that the convergence and accuracy of the solutions are improved for reconstructing natural images.

Dictionary learning methods have also been applied for classification and clustering. Zhang and Li [10] developed a discriminative K-SVD (D-KSVD) method for face recognition. The D-KSVD method is implemented by adding a discriminative term into the objective function of the original K-SVD algorithm. The D-KSVD method outperforms other existing methods such as the SRC algorithm [11]. Ptucha and Savakis [12] proposed a linear extension of graph embedding K-means-based singular value decomposition (LGE-KSVD) method to solve facial and activity recognition problems. LGE-KSVD utilized variants of the LGE to optimize the K-SVD problem. Other dictionary learning methods for classification include label consistent K-SVD [13], discriminative Bayesian dictionary learning [14], and task-driven dictionary learning [15].

These existing dictionary learning methods were developed for either reconstruction or classification. The proposed framework in this paper is for simultaneous reconstruction and classification by optimizing the measurement, basis and classification matrices. Therefore, instead of collecting the complete signal, only a few data points are needed to recover the sparse coefficient vectors. The recovered coefficient vectors can be used to reconstruct and classify the original signal by multiplying the basis matrix and the classification matrix respectively. The measurement matrix is designed to determine the measurement locations in physical experiments, which typically requires that there is only one non-zero entry in each row of the measurement matrix. The consideration of physical constraints and interpretations in measurement conditions is also important in engineering applications.

3. Methodology

The physics-constrained dictionary learning scheme developed in [3] is modified and extended to solve the reconstruction and classification problems simultaneously. The new formulation of the problems is

$$\min_{\Phi,\Psi,\mathbf{C},\mathbf{Y}} \alpha\|\mathbf{S}-\mathbf{\Psi}\mathbf{Y}\|_F^2+\|\mathbf{\Phi}\mathbf{S}-\mathbf{\Phi}\mathbf{\Psi}\mathbf{Y}\|_F^2+\beta\|\mathbf{L}-\mathbf{C}\mathbf{Y}\|_F^2 \tag{1}$$

$$\text{subject to} \quad \mathbf{\Phi}=f(\mathbf{\Psi}) \tag{2}$$

$$\|\boldsymbol{\gamma}_i\|_{\mathbf{0}} \le T_0, \ \ \forall i \tag{3}$$

where F denotes the Frobenius norm, $\mathbf{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2 \dots \boldsymbol{s}_P] \in \mathbb{R}^{N\times P}$ contains P sets of training data and each data set $\boldsymbol{s}_j$ has the length of N. $\boldsymbol{\Psi} \in \mathbb{R}^{N\times W}$ is the basis matrix with $N < W$ and $W \ll P$. $\boldsymbol{\Upsilon} = [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \dots \boldsymbol{\gamma}_P] \in \mathbb{R}^{W\times P}$ contains the sparse coefficients that represent the training data in $\mathbf{S}$ with respect to the basis matrix. $\mathbf{C} \in \mathbb{R}^{N\times W}$ is the classification matrix and $\mathbf{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2 \dots \boldsymbol{l}_P] \in \mathbb{R}^{N\times P}$ are the class labels of the training signals. The index of the non-zero value in $\boldsymbol{l}_i$, such as $\boldsymbol{l}_i = [0, 1, 0 \dots 0]$, indicates the class. For diagnostics, different classes correspond to different machine states. Lagrange multipliers α and β are applied to control the relative contribution of each term in Equation (1). The constraint in Equation (2) indicates the training sequence. That is, in the iterative optimization procedure, the basis matrix $\boldsymbol{\Psi}$ is fixed first, measurement matrix $\boldsymbol{\Phi}$ is optimized based on the fixed $\boldsymbol{\Psi}$. Then the basis and classification matrices are optimized next. The constraint in Equation (3) is the upper limit of the sparsity level, where $\boldsymbol{\gamma}_i$ is the *i*-th column of the coefficient matrix, and T_0 is the target number of non-zero values in the sparse vectors of coefficients. In addition to constraints in Equations (2) and (3), more physical constraints can be added similarly. In this paper, low-resolution images are collected by a low-resolution camera and can be used to reconstruct the high-resolution ones, and machine states can be identified more clearly with the reconstructed high-resolution images. Because the pixels in the low-resolution images correspond a grid, the physical constraint that the optimal locations of the collected pixels need to be close to the grid is considered. Pixel values between grid points which cannot be directly measured are estimated by linearly interpolating the measured pixel values. Other physical constraints can be added similarly.

The learning procedure starts with an initial guess of the basis matrix. Two stages are performed in each iteration. At stage one, the measurement matrix is optimized to determine the locations of measurements with the fixed basis matrix. This can be solved based on the constrained FrameSense algorithm [3]. At stage two, the basis and classification matrices can be optimized with the fixed measurement matrix based on the K-SVD algorithm. The above two optimization steps are repeated until the optimal measurement, classification and basis matrices converge respectively without further improvement.

3.1. Stage One Optimization

At stage one, with the basis matrix $\boldsymbol{\Psi}$ fixed, the measurement matrix $\boldsymbol{\Phi}$ is optimized to determine the locations of collected pixels. Determining the optimal pixel locations from all available ones is often NP-hard if the amount of data collection is large. Therefore, a greedy algorithm called constrained FrameSense [3] is used to determine the near-optimal pixel locations. Given all available pixel locations $\mathcal{N} = \{1, \dots, N\}$, an unsuitable set of pixel locations $\mathcal{T}$ can be iteratively identified as the index of the row in the basis matrix $\boldsymbol{\Psi}$ by solving [16]

$$\max_{\mathcal{T}} \mathrm{F}(\mathcal{T}) = \mathrm{H}(\mathcal{T}) - \mathrm{H}(\boldsymbol{\Psi}_{\mathcal{N}\backslash\mathcal{T}}) \tag{4}$$

where $\mathrm{H}(\boldsymbol{\Psi})$ is the frame potential and represented as $\mathrm{H}(\boldsymbol{\Psi}) = \sum_{i=1}^{N} |\lambda_i|^2$. λ_i is the *i*-th largest eigenvalue of $\boldsymbol{\Psi}^*\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}^*$ is the conjugate transpose of $\boldsymbol{\Psi}$. $\boldsymbol{\Psi}_{\mathcal{N}\backslash\mathcal{T}}$ is a sub-matrix of $\boldsymbol{\Psi}_{\mathcal{N}}$ with rows corresponding to the indices with the unsuitable ones excluded. After determining the unsuitable pixel location $\mathcal{T}$, the new available pixel locations are updated as $\mathcal{N}\backslash\mathcal{T}$. If M pixel values need to be collected, the locations of M pixels are optimized by excluding $(N - M)$ unsuitable pixel locations iteratively. The desirable pixel locations can be identified in the optimized $M \times N$ measurement matrix. There is only one non-zero entry which has the value of 1 in each row of the measurement matrix and the column index of the row indicates the location where the pixel value is collected. The physical constraint that the optimal locations of the collected pixels need to be close to the grid is also incorporated in the constrained FrameSense algorithm. Therefore, in each iteration, if the unsuitable location belongs to the pixel locations in the low-resolution images, this location is remained in the set of the optimized locations. For any pixel value between grid points, it is estimated by linearly interpolating all pixel values at the grid points.

3.2. Stage Two Optimization

The objective function in Equation (1) can be converted to

$$\min_{\mathrm{C},\Psi,\Upsilon} \left\| \begin{pmatrix} \alpha\mathbf{S} \\ \boldsymbol{\Phi}\mathbf{S} \\ \beta\mathbf{L} \end{pmatrix} - \begin{pmatrix} \alpha\boldsymbol{\Psi} \\ \boldsymbol{\Phi}\boldsymbol{\Psi} \\ \beta\mathbf{C} \end{pmatrix} \boldsymbol{\Upsilon} \right\|_F^2 \tag{5}$$

With $\mathbf{X} = [\alpha\mathbf{S} \quad \boldsymbol{\Phi}\mathbf{S} \quad \beta\mathbf{L}]^{\mathbf{T}}$, $\mathbf{Z} = [\alpha\boldsymbol{\Psi} \quad \boldsymbol{\Phi}\boldsymbol{\Psi} \quad \beta\mathbf{C}]^{\mathbf{T}}$ is optimized by solving Equations (5) and (3) with the K-SVD algorithm [6]. The sub-matrix $\mathbf{Z}_1 = [\alpha\boldsymbol{\Psi} \quad \boldsymbol{\Phi}\boldsymbol{\Psi}]^{\mathbf{T}}$ is used to obtain the basis matrix $\boldsymbol{\Psi}$ by solving

$$\mathbf{\Psi} = (\alpha^2\mathbf{I} + \mathbf{\Phi}^{\mathbf{T}}\mathbf{\Phi})^{-\mathbf{1}}[\ \alpha\mathbf{I} \quad \mathbf{\Phi}^{\mathbf{T}}]\mathbf{Z}_1 \tag{6}$$

With the optimized sub-matrix $\mathbf{Z}_2 = \beta\mathbf{C}$, the classification matrix can be obtained as $\mathbf{C} = \mathbf{Z}_2/\beta$. The basis matrix is then normalized as $\mathbf{\Psi}' = \left[\frac{\varphi_1}{\|\varphi_1\|_2}, \frac{\varphi_2}{\|\varphi_2\|_2}, \dots, \frac{\varphi_W}{\|\varphi_W\|_2},\right]$ and the corresponding classification matrix is $\mathbf{C}' = \left[\frac{c_1}{\|\varphi_1\|_2}, \frac{c_2}{\|\varphi_2\|_2}, \dots, \frac{c_W}{\|\varphi_W\|_2},\right]$ where φ_i's and c_i's are columns in the original basis matrix $\mathbf{\Psi}$ and classification matrix $\mathbf{C}$ respectively. The normalized basis matrix $\mathbf{\Psi}'$ and the corresponding classification matrix $\mathbf{C}'$ are used for the reconstruction and classification of testing signals. The proposed physics-constrained dictionary learning algorithm to optimize the measurement, basis, and classification matrices is shown in **Table 1**.

Table 1: Physics-constrained dictionary learning for classification

Initialize the basis matrix $\mathbf{\Psi}$ and m=0;
WHILE $m <$ total number of iterations DO
1. Compute $\mathbf{\Phi}$ based on the constrained FrameSense algorithm
2. With $\mathbf{X} = [\alpha\mathbf{S} \quad \mathbf{\Phi}\mathbf{S} \quad \beta\mathbf{L}]^{\mathbf{T}}$, $\mathbf{Z} = [\alpha\mathbf{\Psi} \quad \mathbf{\Phi}\mathbf{\Psi} \quad \beta\mathbf{C}]^{\mathbf{T}}$ and $\mathbf{\Upsilon}$ are updated by solving Equations (5) and (3) with K-SVD method.
3. Update $\mathbf{\Psi}$ with Equation (6) and $\mathbf{C} = \mathbf{Z}_2/\beta$
4. Normalize $\mathbf{\Psi}$ as $\mathbf{\Psi}'$ and obtain the corresponding $\mathbf{C}'$
5. $m = m + 1$
END WHILE

4. Experiments

In the SLM process, shape and size of melt pool need to be monitored and controlled. They are indicators of the temperature gradients and cooling rates, which affect the formation of grain structures and defects during the rapid solidification. The size of the laser beam spot is an important factor that affects the geometry of the melt pool. Powder spatter during the manufacturing process is another major cause of defect formation. Spatters also cause contamination in the powder bed. Therefore, monitoring the laser beam spot and powder spatter in the SLM process is critical to control the quality of the solid build.

(a) (b) (c)

Figure 1: The high-resolution optical images collected when (a) the laser is off at the turning point, (b) the laser is on and there is a small amount of powder spatter, and (c) the laser is on and there is a large amount of powder spatter

A high-speed optical camera was used to monitor the laser beam spot and powder spatter. A total of 1700 images were collected. To demonstrate the proposed physics-constrained dictionary learning method, the original images with the size of 200×200 pixels are rescaled to low-resolution ones with the size of 25×25 pixels. 1400 images are used for the training process and the remaining 300 images are used for testing. The collected images are classified into three categories, and one example in each category is shown in **Figure 1**. The first category includes images collected when the laser is turned off at turning points as shown in **Figure 1** (a). The second category includes images collected in the normal manufacturing process, where the laser is on and there is a small amount of powder spatter as shown in **Figure 1** (b). The third category includes images collected in the defective manufacturing process, where the laser is on and there is a large amount of powder spatter as shown in **Figure 1** (c). Our goal is to identify the system's state based on low-resolution images.

The pixel locations optimized with the physics-constrained dictionary learning method for desired $M = 121$ pixel values are shown in **Figure 2** (a), which include 81 that are directly measurable in the regular grid and 40 that are not directly measurable. The pixel values at the locations marked with the square symbol which cannot be measured directly are estimated by linearly interpolating all pixel values marked with the circle symbol, which are obtained directly from the low-resolution images. With the optimized basis matrix and the pixel values at the locations in **Figure 2** (a), the coefficient vector can be recovered with the OMP algorithm [17]. Then high-resolution images can be obtained as the linear combinations of the basis matrix and the recovered coefficient vectors. The original images and reconstructed images by CS and physics-constrained dictionary learning are in **Figure 2** (b) and (c) respectively. In order to show the advantage of the proposed method for the image reconstruction compared to the traditional techniques such as the simple linear interpolation, the images that are reconstructed by simple linear interpolation of the measured pixel values in the low-resolution images are also shown in **Figure 2** (d). Compared to the results from the simple linear interpolation, the reconstructed high-resolution images with CS and the physics-constrained dictionary learning method can be used to identify the laser beam spot and powder spatter more clearly. The class labels of the testing images can be obtained by the linear combinations of the optimized classification matrix and the recovered coefficient vectors, which are used to identify different machine states in **Figure 1**. The classification error for all testing images is 16%, which is computed as $e_c = N_I/N_T \times 100\%$ where N_I is the number of incorrect class labels and N_T is the number of total class labels.

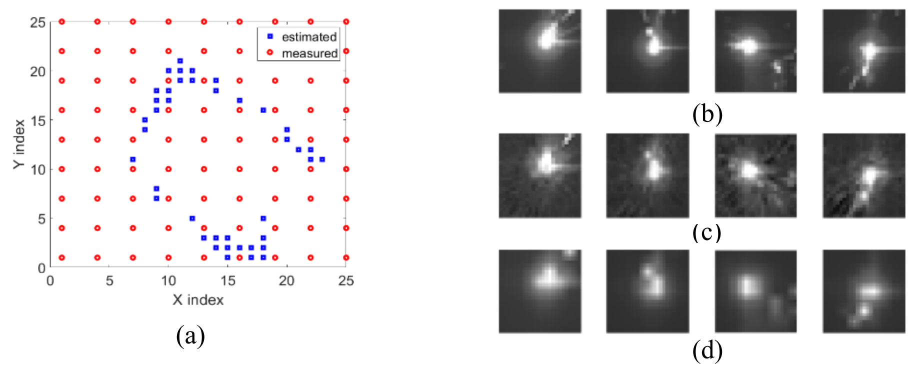

Figure 2: (a) Optimized pixel locations, (b) original images, (c) reconstructed images by CS and physics-constrained dictionary learning, and (d) reconstructed images by interpolation.

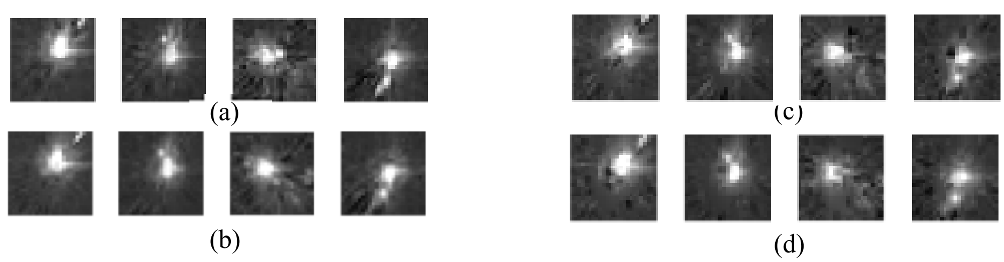

Figure 3: Reconstructed images by CS and physics-constrained dictionary learning when (a) $M = 81$, (b) $M = 121$, (c) $M = 150$, and (d) $M = 180$.

Sensitivity analysis is performed with different values of M as shown in **Figure 3**. It is found that as fewer pixel values are collected, the reconstructed images contain more noises. However, when more pixel values are collected, the reconstructed images can also be more blurred, because more interpolation errors are included in the reconstruction results when more pixel values between grid points are estimated and used. Therefore, the trade-off between noise and interpolation errors needs to be considered when the desired number of collected pixel values is determined. The classification errors are similar with different values of M because the features in the image such as the laser spot and power spatter can still be identified from the reconstructed images of lower quality in **Figure 3**.

5. Conclusion

In this paper, a new physics-constrained dictionary learning method is proposed to solve reconstruction and classification problems by optimizing measurement, basis, and classification matrices simultaneously. The proposed method is demonstrated to monitor the SLM process with optical images. With low-resolution images, high-resolution ones can be reconstructed to identify different machine states. Therefore, the required memory usage for data collection can be significantly reduced in process monitoring.

The major challenge of the proposed physics-constrained dictionary learning comes from the training algorithm that is based on the K-SVD. The K-SVD can only find the local optima. The performance of reconstruction depends on the choices of the initial basis matrix and the reconstruction algorithm. In this work, the initial basis matrix is generated by randomly selecting a few columns of the training dataset. However, selected columns may not represent the characteristics of all categories since the training dataset consists of signals from different classes. Therefore, an initialization method needs to be developed to further improve the classification accuracy. Here, the physics-constrained dictionary learning method is used to capture the spatial correlations in the optical images. The physical constraint that the distribution of measured pixels needs to be close to the grid is applied. In future work, the method will be used for higher dimensional data such as videos. The temporal correlations between the successive images will also be considered and used as additional physical constraints. In large-scale sensor networks, the physical constraints can be designed based on the limitations of communication between sensors and the coverage of areas.

Acknowledgements

This work was supported in part by the George W. Woodruff Faculty Fellowship and National Science Foundation under grant CMMI-1663227.

References

[1] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Transactions on Information Theory,* vol. 52, no. 12, pp. 5406-5425, Dec 2006.
[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory,* vol. 52, no. 4, pp. 1289-1306, Apr 2006.
[3] Y. Lu and Y. Wang, "A Physics-Constrained Dictionary Learning Approach for Compression of Vibration Signals," *Mechanical Systems and Signal Processing,* vol.153, pp.107434, 2021.
[4] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T. W. Lee, and T. J. Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural Computation,* vol. 15, no. 2, pp. 349-396, Feb 2003.
[5] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," *Icassp '99: 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings Vols I-Vi,* pp. 2443-2446, 1999.
[6] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing,* vol. 54, no. 11, pp. 4311-4322, Nov 2006.
[7] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online Learning for Matrix Factorization and Sparse Coding," *Journal of Machine Learning Research,* vol. 11, pp. 19-60, Jan 2010.
[8] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to Sense Sparse Signals: Simultaneous Sensing Matrix and Sparsifying Dictionary Optimization," *IEEE Transactions on Image Processing,* vol. 18, no. 7, pp. 1395-1408, Jul 2009.
[9] H. Bai, G. Li, S. Li, Q. W. Li, Q. R. Jiang, and L. P. Chang, "Alternating Optimization of Sensing Matrix and Sparsifying Dictionary for Compressed Sensing," *IEEE Transactions on Signal Processing,* vol. 63, no. 6, pp. 1581-1594, Mar 15 2015.
[10] Q. A. Zhang and B. X. Li, "Discriminative K-SVD for Dictionary Learning in Face Recognition," *2010 IEEE Conference on Computer Vision and Pattern Recognition (Cvpr),* pp. 2691-2698, 2010.
[11] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, no. 2, pp. 210-227, Feb 2009.
[12] R. Ptucha and A. E. Savakis, "LGE-KSVD: Robust Sparse Representation Classification," *IEEE Transactions on Image Processing,* vol. 23, no. 4, pp. 1737-1750, Apr 2014.
[13] Z. L. Jiang, Z. Lin, and L. S. Davis, "Label Consistent K-SVD: Learning a Discriminative Dictionary for Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 35, no. 11, pp. 2651-2664, Nov 2013.
[14] N. Akhtar, F. Shafait, and A. Mian, "Discriminative Bayesian Dictionary Learning for Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 38, no. 12, pp. 2374-2388, Dec 2016.
[15] S. Bahrampour, N. M. Nasrabadi, A. Ray, and W. K. Jenkins, "Multimodal Task-Driven Dictionary Learning for Image Classification," *IEEE Transactions on Image Processing,* vol. 25, no. 1, pp. 24-38, Jan 2016.
[16] J. Ranieri, A. Chebira, and M. Vetterli, "Near-Optimal Sensor Placement for Linear Inverse Problems," *IEEE Transactions on Signal Processing,* vol. 62, no. 5, pp. 1135-1146, Mar 2014.
[17] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory,* vol. 53, no. 12, pp. 4655-4666, Dec 2007.