Learning Sparsifying Transforms For Compressively Sampled Biomedical Imaging Modalities

By

Shahid Ikram

Reg. No. 61-FET/PHDEE/F13

A dissertation submitted to I.I.U. in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering
Faculty of Engineering and Technology
INTERNATIONAL ISLAMIC UNIVERSITY
ISLAMABAD

2020

I Magnetic resonance imaging

II Biomedical imaging

DEDICATED TO

My Teachers,

Parents and Family

CERTIFICATE OF APPROVAL

Title of Thesis: **Learning Sparsifying Transforms For Compressively Sampled Biomedical Imaging Modalities**

Name of Student: Shahid Ikram

Registration No.: 61-FET/PhDEE/F13

This thesis is accepted by the Department of Electrical Engineering, Faculty of Engineering and Technology, International Islamic University, Islamabad, in partial fulfillment of the requirements for the PhD degree in Electronics Engineering.

VIVA VOCE COMMITTEE

Supervisor

Dr. Jawad Ali Shah

Assistant Professor, DEE, FET, IIUI

Co-Supervisor

Prof. Dr. Ijaz Mansoor Qureshi

Professor, DEE, Air University, Islamabad

Internal Examiner

Dr. Suheel Abdullah Malik

Associate Professor, DEE, FET, IIUI

External Examiner 1

Dr. Muhammad Usman

Nescom, Islamabad

External Examiner 2

Dr. Muhammad Mukhtar Talha

Principal Scientific Officer, KRL

Chairman DEE

Dr. Suheel Abdullah Malik

Associate Professor, DEE, FET, IIUI

Dean FET

Prof. Dr. Muhammad Amir

Professor,

Department of Electrical Engineering, FET, IIUI

ABSTRACT

The objective of this dissertation is to achieve better image quality (IQ) by using the novel compressed sensing (CS) dictionary learning methods that can efficiently recover the medical image(s) from under-sampled (partial Fourier encoded) image with reduced number of measurements.

The application of CS to biomedical imaging largely depends on the use of various sparsifying transforms, such as wavelets (WT), curvelets or total variation (TV) to recover medical resonance (MR) images.

Images reconstruction from CS consists of two steps i.e. dictionary learning and sparse coding. In this thesis, two novel techniques are presented to recover highly under-sampled MR image with better IQ.

In first technique, simultaneous code word optimization (SimCO) patch-based dictionary learning is chosen that updates the atoms simultaneously whereas focal underdetermined system solver (FOCUSS) is used on behalf of sparse representation because of soft constraint on sparseness of an image. Combining SimCO and FOCUSS, it is termed as new scheme called SiFo. The proposed reconstruction scheme initially learns the dictionary and then uses it to eliminate under-sampling reconstruction noise.

Other novel technique is block dictionary learning. The block structure of a dictionary can be learned by exploiting the latent structure of the desired signals. Such type of block dictionary leads to block sparse representation of the given medical signals which has shown better results for reconstruction of the medical images. In this method, we suggest a framework for MRI reconstruction based upon block dictionary. The proposed technique develops automatic detection of underlying block structure of MR

images given maximum block sizes. This is done by iteratively alternating between updating the block structure of the dictionary and block-sparse representation of the MR images.

Based on dictionary learning, this work has shown superior reconstruction of images in the empirical results for noisy and noiseless cases. The performance is validated by using different sampling masks and k-space under sampling ratios. Fast convergence with more accurate reconstruction at high under-sampling is achieved by this scheme with robust performance in a noisy environment.

On the basis of diverse performance parameters, it is evident that the proposed designed techniques outperform the leading state of the art dictionary learning based MRI (DLMRI) reconstruction method.

LIST OF PUBLICATIONS AND SUBMISSIONS

[1] **Ikram, S.**, Zubair, S., Shah, J. A., Qureshi, I. M., Wahid, A., & Khan, A. U. (2019). Enhancing MR Image Reconstruction Using Block Dictionary Learning. IEEE Access, 7, 158434-158444.

[2] **Ikram, S.**, Shah, J. A., Zubair, S., Qureshi, I. M., & Bilal, M. (2019). Improved reconstruction of MR scanned images by using a dictionary learning scheme. Sensors, 19(8), 1918.

[3] Haider, H., Shah, J. A., **Ikram, S.**, & Abd Latif, I. (2017, September). Sparse signal recovery from compressed measurements using hybrid particle swarm optimization. In 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA) (pp. 429-433). IEEE.

ACKNOWLEDGEMENTS

In the name of Allah (SubhanahuWaTa'ala), who is the most kind and the generous. I am grateful to ALLAH SWT for granting me strength, stamina and direction throughout this work and for thoughts, which originated in my mind to complete this research work. Indeed, there is nothing that can happen without HIS will. Peace and blessings of Allah be upon His last Prophet Muhammad (Sallulah-o-Alaihihe-Wassalam) and all his Sahaba (Razi-Allah-o-Anhum), who sacrificed their lives and wealth for spreading Islam on this globe

Grateful to my supervisor Dr. Jawad Ali Shah, whose encouragement, concepts and efforts create an opportunity for me to complete my advanced educations. He appeared as a role model for me in learning, research and other parts of life. I would wish to express gratitude to my co-supervisor Prof. Dr. Ijaz Mansoor Qureshi for his continuous support during my research work.

I would also love to pay special thanks to my virtual supervisor Dr. Syed Zubair for his enormous guidance and the many extensive, appealing, and rewarding debates, remarks, and recommendations throughout this work. In every uncertain problem, he came up with the solutions and kept me encouraged to tackle all the challenges. It was difficult to complete my thesis without his assistance.

I extend my sincere acknowledgments and gratitude to my associates Dr. Aqdas Naveed Malik, Dr. Muhammad Amir, Dr. Suheel Abdullah, Dr. Abdul Jalil, Dr. Hammad Omer, Dr. Adnan Umar Khan, Dr. Muhamad Bilal, Dr. Muhammad Waseem Khan, Dr. Naveed Ishtiaq, Dr. Nouman, Dr. Zeeshan Aslam, Engr. Haris Anees, Engr. Abdul Wahid Tareen, Engr. Sharjeel Abid Butt and, Engr.M Muzzammil for their never ending assistance and for their valuable debates during my advanced studies. I would wish to recognize the kindness and curtesy of International Islamic University Islamabad (IIUI) Pakistan for allowing me complete fee waiver during the PhD educations. I would like to express my *appreciation to British Malaysian Institute (BMI) University of Kuala Lumpur, Malaysia for their financial funding to complete my research work.*

I am surely thankful to my parents (Late), sisters and brothers for their affection and sentiments throughout my life. I am also very thankful to my devoted wife for her patience, unending encouragement and prayers during every stage of my PhD studies. Finally, I am delighted to express my gratitude to my sweet kids Sheza, Zainab, Saad and Ibrahim, whose sinless and innocent gestures were source of inspiration for me.

(Shahid Ikram)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADMM Alternating Direction Method of Multipliers

ARS American Radiology Services

BLKSVD Block-KSVD

BOMP Block – OMP or (Block Orthogonal Matching Pursuit)

BP Basic Pursuit

BBP Block Basic Pursuit

BSL Bayesian Sparse Learning

CT Computed Tomography

CoSaMP Compressive Sampling Matching Pursuit

CNN Convolution Neural Networks

CS Compressed (or compressive) Sensing (or sampling)

CSMRI Compressed Sensing MRI

dB Decibel

DLMRI Dictionary Learning based MRI

DWT Discrete Wavelet Transform

FFT Fast Fourier Transform

DCT Discrete Cosine Transform

DL Dictionary Learning

DFT Discrete Fourier Transform

FISTA Fast Iterative Shrinkage Thresholding

FOV Field of View

FOCUSS FOCal Underdetermined System Solver

FT Fourier Transform

fMRI Functional Magnetic Resonance Imaging

HPSO Hybrid particle swarm optimization

IFFT Inverse Fast Fourier Transform

ILS Iterative Least-Squares

IQ Image Quality

ISA Iterative-shrinkage algorithm

K-SVD K-mean Singular Value Decomposition

HFEN High Frequency Error Norm

LASSO Least Absolute Shrinkage and Selection Operator.

LOG Laplacian of Gaussian

NP Non-deterministic Polynomial

PCD Parallel Coordinate Descent

PSO Particle Swarm Optimization

PET Positron Emission Tomography

RLS Recursive Least Squares

RMS Root Mean Square

ROMP Regularized Orthogonal Matching Pursuit

RIP Restricted Isometry Property

SSIM Structural Similarity Index

SNR Signal-to-Noise Ratio

SC Sparse coding

SIFO Simultaneous code word optimization-FOCal Underdetermined System Solver

SimCo Simultaneous Code word Optimization

SSF Separable Surrogate Functionals

SVD Singular Value Decomposition

MR Magnetic Resonance

MOD Method of Optimization Direction

MWI Microwave Imaging

MRI Magnetic Resonance Imaging

NP Non-deterministic Polynomial-time

MSE Mean Square Error

OMP Orthogonal Matching Pursuit

PSNR Peak Signal-to-Noise Ratio

TV Total Variation

ZF Zero-Filling

Chapter 1

Introduction

1.1 Introduction

Medical imaging techniques are playing a vital role in medical diagnostic to study the insight of human body such as X-rays, computed tomography (CT), microwave imaging (MWI) and magnetic resonance imaging (MRI)

MRI is one of the important tools to generate anatomical images of the body without the use of damaging and harmful radiation, unlike CT scan or X-rays. A strong magnetic field and radio waves are used in MRI to produce detailed images of the organs and soft tissues within the body. These days, MRI is considered as the main source of reproducible diagnostic medical information because of noninvasive technique and accurate visualization of the anatomical skeleton. It is essential to take a higher number of measurements to cover a huge dynamic series of tissue parameter in related medical application for reconstruction of the image with clear features.

Latest techniques in biomedical imaging have played a very important role for clinical applications in helping the doctors for better diagnostic and treatment. Main research objectives are as follows:

1. To use sparsifying transforms instead of the conventional orthogonal wavelets, discrete cosine or total variation for improving the reconstruction accuracy in compressed sensing.
2. To develop new algorithms that can be used to improves the speed of MRI acquisition by including sparsifying transforms during the reconstruction.
3. To investigate and modify recently developed algorithms used for sparsifying transforms and sparse coding to attain the better results in terms of IQ.
4. To learn sparsifying transform for patch based image instead of whole image to cater local features of image more accurately.
5. To ease the approximate sparse representation, block sparsity concept can be introduced in sparsifying transform by exploiting the latent structure of desired signal.
6. To reconstruct the low dimension signals/objects by using sparsifying transform techniques e.g. synthetic-aperture radar (SAR) or Tele diagnosis

For better resolutions of recovered medical images, it is essential to explore new tools and techniques that can sparsely approximate and represent the biomedical images. That will be useful for extracting the fruitful and desirable clinical information efficiently and hence decreasing the diagnostic time.

To accomplish the better image quality (IQ) for image reconstruction, we can achieve it by considering and adopting better sparsifying transforms for CS in MRI. There is a remarkable potential available, for recovery of images with high accuracy, to study the number of transforms. Since sparsifying transforms have delivered the improvement and advancement for the CS solution in MRI, Therefore, investigating and modification in different algorithms for sparsifying transforms and sparse coding in recently developed techniques; one can attain the better results in IQ. Patch based and block based dictionary learning lead for development in new algorithms for recovery of images recently.

Main problem in MRI is the time acquisition to get large numbers of image samples for a better reconstruction of the image. A patient must wait for a long period of time in MRI scanner. The problem of long acquisition time in a scanner can be reduced to make development either on hardware side or software side. The changes in software side can be implemented more easily than hardware side through efficient algorithms. These algorithms mainly depend on CS technique.

The CS theory in biomedical imaging application has become popular as it permits a rationally almost exact reconstruction of images from far fewer measurements. For biomedical imaging, CS can increase the imaging speed and consequently decrease the radiation dose. Modern theory of CS [1]–[8][9]–[18] supports the reconstruction of signals accurately by means of fewer information as compared to the set of unknowns, or required by conventional Nyquist sampling. It is only possible, on condition that the underlying signal is nearly sparse in any transform domain. The consequence of this improvement is that the reconstruction technique is nonlinear. In recent times, CS theory has been used in MRI [9], [19]–[23] showing good quality of recovery of signal from a reduced amount of information (values).

The superiority of CS reconstruction methods generally relies on the use of various sparsifying transforms also called dictionaries such as wavelet or total variation (TV) to recover MRI, CT [24] and other biomedical images from the subsampled data by exploiting the sparsity of these images in the transform domain or dictionary. Non-adaptive global sparsifying transform, known as compressed sensing MRI (CSMRI) is restricted in conventional MR images to 2-4 folds under-sampling [14], [16], [17], [25]. Numerous unwanted artifacts and loss of features were observed during the reconstruction of images with CSMRI technique and non-adaptive sparsifying transform like wavelet and TV etc. [21], [26].

Initially, predefined dictionaries or sparsifying transforms known as non-adaptive sparsifying transform were used to reconstruct the medical images. In this thesis, we develop such formulations which aim to learn sparsifying transforms from data. Adaptive learning sparsifying transforms can better sparsify the images because these are trained from the particular class of images [27], [28]. Different artifacts and aliasing effect come into play on the edges of reconstructed images, when using under-sampling the data from k-space. So noise artifacts due to under-sampling are one of the main challenges to reconstruct the MR images. Patch based dictionary learning has got the popularity, because, it has a tendency to effectively apprehend the local image features and have a potential to eliminate the aliasing artifact without compromising on resolution. An adaptive patch based dictionary is learned from small number of k-space samples which produces promising reconstruction.

Learning sparsifying transform involves a two-step process, i.e. the dictionary learning from training data and sparse representation. Popular (algorithms) techniques for dictionary learning are method of optimal directions (MOD) [27], K-times singular value decomposition (KSVD) [28], alternating direction method of multipliers (ADMM) [29], iterative least square (ILS) [30], recursive least square (RLS) [31] and simultaneous coded optimization SimCO [32], [33] whereas orthogonal matching pursuit (OMP) [34], [35], basic pursuit (BP) [36],[37], [38] block basic pursuit (BBP) [39],[40],[37], block orthogonal matching pursuit (BOMP) [25]–[27],[28], least absolute shrinkage and selection operator (LASSO), [44], [45], and (FOCUSS) [46] techniques are used for sparse coding. These learning algorithms face the challenge in finding a dictionary to express optimal sparse representations for the image recovery.

A well-known algorithm K-SVD is used extensively for dictionary learning along with OMP for sparse coding to reconstruct the image. The K-SVD updates the columns (called atoms) sequentially one by one and takes time to update all the atoms. Hence it is computationally expensive. Increase of singular vectors is also a problem in this technique. Although, K-SVD performs better in capturing a reference dictionary, but its de-noising performance is comparatively limited. Conventional sparse coding technique OMP causes to produce the artifacts at the edges due to the hard constraint. The hard sparseness constraint may not be very useful for a medical imaging application which produces artifacts on a high frequency. Most of the researchers use the KSVD technique for dictionary learning and apply OMP for sparse coding. One of the leading technique for reconstruction of MR images is DLMRI by Saiprasad et al [26]. His framework was based on K-SVD for learning sparsifying transform with OMP method for sparse coding.

Signal or image can be extracted from union of a small number of subspaces [28],[37], [40]. Dictionary atoms are categorized into underlying subspaces in such kind of signals which lead to sparse representation for block sparse structure [43],[47]. There are many methods, such as BBP [27], [41],[37], BOMP [42],[39] and group Lasso [45], [48] that have been suggested to get the benefit of this framework in reconstruction of the block sparse representation. Block structured dictionaries are learned to utilize any embedded structure in order to produce a more effective sparse representation. In many applications, signals can be sorted into blocks and approximated into block sparse. This leads the concept of block dictionary learning. Block-structured dictionary method also involves into two steps; update the block dictionary and then finding the sparse coefficients according to block structure in the dictionary. Algorithm BOMP picks the dictionary blocks sequentially which are best suited to the applied signals. Block-SVD (BLKSVD) technique updates the atoms in blocks and corresponding non-zero coefficients simultaneously.

A lot of research work is being carried out on deep learning, particularly in convolutional neural networks (CNNs) for MRI reconstruction [49],[50]–[53] but the main bottle neck is the availability of high computation resources and large amount of data. In our proposed work, we learn the dictionary for a single image. So single image (slice) cannot be used for training a deep learning network. Hence, we have restricted our work to a dictionary learning based method instead of deep learning (CNN).

To address the above-mentioned issues: like hard constrained in sparse coding, update the dictionary atoms one at a time, increase of singular vectors, limited performance on de-noising and removal of artifacts and aliasing effect and fast convergence; the proposed framework and methods are applied on the recovery of biomedical imaging modalities using Shepp-Logan phantom as well as original MR images. The performance of the proposed algorithm is validated with various under sampling factors, for noiseless and noisy cases. The under-sampling is directly applied on k-space (fully-sampled) MR data set. Axial T2-weighted reference images of the brain are used as MR images, taken from vivo MR scans of size 512×512 from American Radiology Services (ARS) [26],[54]–[56]. Efficiency of the proposed methods have been evaluated using recovered images qualitatively and in terms of various performance parameters such as high frequency error norm (HFEN), peak signal to noise ratio (PSNR), Structural similarity index measurement (SSIM), correlation and sharpness index (SI).

1.2 Main Contribution

Primary contribution of this thesis is the improved recovery of compressively sampled biomedical imaging modalities such as MR (brain) image based on learning sparsifying transforms schemes. Contributions of this dissertation are discussed as follows:

We have proposed adaptive patch-based dictionary learning by introducing a hybrid algorithm of SimCO along with FOCUSS (we termed it as SiFo) for MR images. Since dictionary learning involves two step process i.e. sparse coding and dictionary update, so in this novel hybrid SiFo technique we have used FOCUSS as a sparse coding technique because of its soft constraints, which is considered to be better for sparse representation of medical images. The pruning process of FOCUSS has a tendency to suppress the noise during the reconstruction of image because aliasing artifacts and noises are usually isolated. This is the main reason for our proposed scheme, which performs better particularly in noisy cases. While SimCO is used for dictionary updates. SimCO method simultaneously updates an arbitrary number of atoms instead of atom by atom such as in KSVD method. The problem of increased singular vectors in K-SVD is minimized by using a regularized version of SimCO. In both noisy and noiseless cases, our technique provides superior reconstruction of images. The performance is confirmed by the collection of various sampling trials and k-space under-sampling aspects. Fast convergence with accurate reconstruction at high under sampling rates is achieved by this proposed algorithm.

The other major contribution of this thesis is to introduce block dictionary learning called "Block K-SVD" for effectively recover the real MR image from its under-sampled k-space data. The motivation for considering block-sparse signals is represented in a twofold aspect. First, in numerous examples the nonzero values of sparse vectors incline to group in blocks; a number of illustrations and samples are specified in [57]. Second, the situation is revealed in [57] that sampling problems over unions of subspaces can be transformed into block-sparse reconstruction problems. To exploit the property of MRI data to represent it into cluster or structured form, we have presented the block dictionary learning concept into recovery of medical image modalities first time. The proposed method is an extension of K-SVD where blocks of atoms are learned and updated iteratively, instead of updating single atom on each iteration. Sparse coding algorithm BOMP chooses the dictionary slabs (blocks) sequentially which are best suitable to the input signals. BLKSVD technique updates the atoms in blocks and corresponding non-zero coefficients simultaneously.

Key points of novelty of this thesis for dictionary learning for reconstruction of CS MR image(s) are as follows:

(1) Introducing of hybrid algorithm of SimCO and FOCUSS, called as SiFo, for reconstruction of under-sampled MR image.

(2) MRI reconstruction problem has been formulated as learning block structure of the MRI that can lead to block dictionary learning. Block dictionary learning helps in finding robust MRI reconstruction.

(3) Based upon block dictionary learning, MRI representation is formulated in terms of block sparsity instead of simply sparsity that leads to better reconstruction.

(4) The block structure exploitation of the MRI for block sparse representation is tested for different sampling schemes to validate MRI reconstruction based upon block sparsity.

(5) We have initialized the dictionary by extracting left singular vectors from the training data for block dictionary learning in our framework and then normalized each atom of the dictionary. We have observed that by initializing dictionary in this form, convergence becomes faster as compared with randomly initialized dictionary

1.3 Thesis organization

Chapter 2 describes literature review for sparse model, CS theory, dictionary learning and different models of sparsifying transform to represent a MR image as a linear combination form and its sparse representation. Different compression technique CSMRI and DLMRI are described in detail with its advantages along with some shortcomings. Dictionary learning and its challenges block dictionary learning and performance parameters are also discussed in this chapter.

Chapter 3 introduces adaptive patch-based diotionary learning framework in which hybrid algorithm of SimCO and FOCUSS is introduced for the recovery of MR images. Since proposed framework involves into two step procedure i.e. SimCo is used to update the dictionary atoms and FOCUSS is used as sparse coding step. In experiments and discussion part, we have evaluated our results with leading stat of the art technique i.e. DLMRI which had already leads over Lustig technique LDP [21]and Zero filling.

Chapter 4 explains the recovery of MR image through block dictionary learning scheme .In this proposed method, BOMP is used as sparse coding to select the dictionary blocks sequentially that are best suited to the input signals. The key characteristic of BLKSVD (block dictionary) is to update the atoms in blocks and corresponding non-zero coefficients simultaneously.

Chapter 5 describes the conclusion along with the future work. Adaptive patch-based block-structured dictionary learning and hybridization of two algorithms (called SiFo) framework have been presented for reconstruction of MR images using the proposed technique. The proposed methods have shown improved performance over other dictionary learning based methods such as DLMRI, for both noisy and noiseless cases. The performance is validated by using a diversity of sampling trails and k-space under sampling ratios.

Chapter 2

Literature Review

Application of CS to MRI provides an opportunity in term of drastically decreased scan time, with compensations for patients and medical management economics. MRI signals are inherently compressible due to their acquisition in transform domain (k-Space/Fourier) and hence are naturally sparse. CS exploits the sparsity of MR images to support the reconstruction from under-sampled (Fourier) k-sample data with high accuracy. Different methods and techniques based on CS theory have been developed recently for recovery of images. These techniques are mainly categorized into non-adaptive and adaptive dictionaries, based on applications. This chapter describes in detail about the recently developed methods for sparse representation of signals and their reconstruction.

2.1 Medical Imaging Techniques

Now days, researchers have focused on sparse representation of signals or medical images because of its popularity. Main goal of sparse representation is to approximate a natural signal or medical image by representing it in linear combination with few significant (non-zero) coefficients. These sparse signals are expressed in some fixed collection (called dictionary) or analytical transform domain such as TV, wavelets [58] and discrete cosine transform (DCT) to recover MRI, CT and other biomedical images from subsampled data by exploiting the sparsity of these images in the transform domain.

Instead of using the analytical transforms, the idea of learning the sparsifying transforms for a specific biomedical imaging modality has received more attention because of its superiority of image representation over the analytical transforms. In recent past, the sparsity based techniques are becoming popular in various applications related to the biomedical imaging. These techniques mainly utilize the image sparsity or its patches in a sparsifying domain to de-noise, restore or recover the images [59]–[61].

Dictionary learning has the remarkable capabilities to expose a priori unknown data of certain kinds of signals taken by different equipment and machine. Examples are medical signals, such as MRI, functional MRI (fMRI), electroencephalogram (EEG), ultrasound tomography (UST) and electrocardiography (ECG) where signals are produced by various physical causes. However, the representation, de-noising, and signal analysis are executed in the correct signal subspace, in a way that observed signals can be recognized by their original physical causes. Adjusting instrument in ECG signals helps ventricular erasure and atrial demonstrating in the ECG of patients experiencing sporadic heartbeat called atrial fibrillation (AF) [62]. Over-complete dictionaries learned, from MRI data (brain and breast tissues), have delivered a tremendous sparse representation and provided better breast tissue images acquired by the UST scanner[63]. Dictionary learning methods involve in the investigation of other signals such as neural signals obtained by EEG, multi-electrode arrays, or two-photon microscopy. This chapter serves as literature review for the dictionary learning and conducted research.

2.2 Sparse Model

Approximate representation of any signal as a linear combination form can be done by just few elements (non-zero coefficients) from a known basis or dictionary is called a sparse signal and also known as linear sparse model. Such kind of model offers a mathematical scheme for apprehending the fact that in many cases, high dimensional signals hold comparatively fewer information as related to their ambient dimension. To understand the theory of sparsity of signal in mathematics, lets defined as: Suppose we have a discrete time signal $\boldsymbol{x} \in \mathbb{R}^q$ to be expressed as a set of support vectors $\{\boldsymbol{\Psi}_i\}_{i=1}^{i=q}$ (called orthogonal basis) as follows:

$$\boldsymbol{x} = \sum_i^q c_i\, \boldsymbol{\Psi}_i = \boldsymbol{\Psi c} \tag{2.1}$$

Where c_i represent the coefficient of the sequence of $\boldsymbol{x}$.Mathematically, signal $\boldsymbol{x}$ is k - sparse when it contain k nonzeros values i.e. $\|\boldsymbol{x}\|_0 \leq$ k. Typically, the signal $\boldsymbol{x}$ is not sparse themselves but may sparse with few basis vectors in $\boldsymbol{\Phi} \in \mathbb{R}^{f \times q}$. In such scenario, $\boldsymbol{x}$ is considered as k sparse; with understanding $\boldsymbol{x}$ is the sparse code. Expressing $\boldsymbol{x}$ as:

$$\boldsymbol{y} = \boldsymbol{\Phi x} \quad \text{s.t.}\ \|\boldsymbol{x}\|_0 \leq \text{k} \tag{2.2}$$

Figure 2.1 shows the simple linear sparse model for representation a signal $\boldsymbol{x}$ with an over complete dictionary $\boldsymbol{\Phi}$ such that $f < q$.

In CS theory, the given image, in encoded mechanism, with sensing or measurement matrix $\boldsymbol{\Phi}$ can be expressed as a linear transformation ($\boldsymbol{\Phi}: \mathbb{R}^q \rightarrow \mathbb{R}^f$) of a length q vector $\boldsymbol{x}$ to a length f measurement vector y with $f \ll q$:

$$\mathbf{y} = \boldsymbol{\Phi x} \quad \text{or} \quad y_j = \sum_k^q \varphi_{jk}\, x_k = \langle \varphi_j, x_k \rangle \quad \text{where } 1 \leq j \leq f \tag{2.3}$$

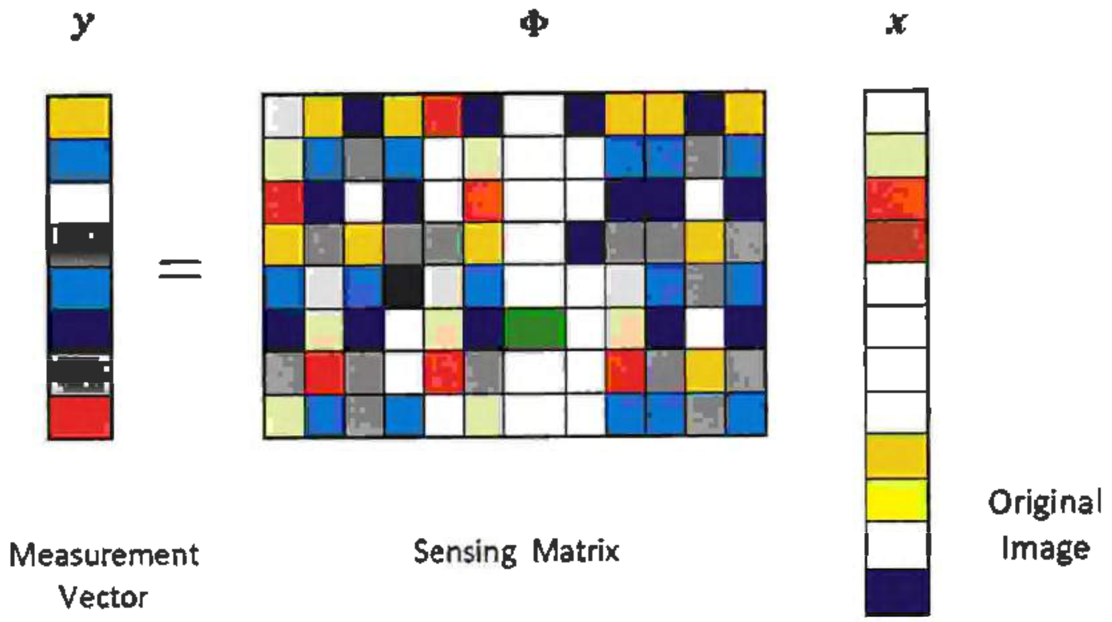


Figure 2.1 *Linear sparse model representation*

Observation or measurement vector **y** is obtained from the non-adaptive linear projection (or transformation) of the original image $\boldsymbol{x}$. Biomedical MR images are encoded into Fourier domain (called k-space) by the scanner, instead of pixel domain. That is why; original image $\boldsymbol{x}$ in Equation (2.3) is represented as measurement vector **y** with respect to encoded (measurement) matrix $\boldsymbol{\Phi}$. When $(q = f)$, measurement matrix become square and invertible, then transformation is reversible which mean that input image or signal can be exactly recovered from the output i.e. $(\hat{\boldsymbol{x}} = \boldsymbol{\Phi}^{-1}\mathbf{y})$. However, in the case $(f \ll q)$, where the numbers of rows are lesser than columns, which leads to an under-determined system of linear equations. In context of linear algebra, it seems to be impossible to find the exact solution for such a case where transformation matrix is rectangular or fat matrix. But CS based framework helps to find the solution in such cases. In CS theory, it is equivalent to the obtaining small amount of measurements than the original image dimension however in the biomedical imaging application e.g. CT and MRI, which directly relates to a decrease in the scan time and a reduced amount of exposure to the electromagnetic radiations (EMR) respectively [20],[64]. Sensing matrix Φ has a null space in an under determined system, which indicates that various vectors outcome with same values or measurements after the transformation. Therefore, the recovery process become ill-condition because there are many solutions. In MRI case, recovery of linear image from its partial Fourier data faces the aliasing artifacts due to the violation of Nyquist sampling criterion .Figure (2.2) shows aliasing artifacts during linear reconstruction of under-sampled brain MR image.

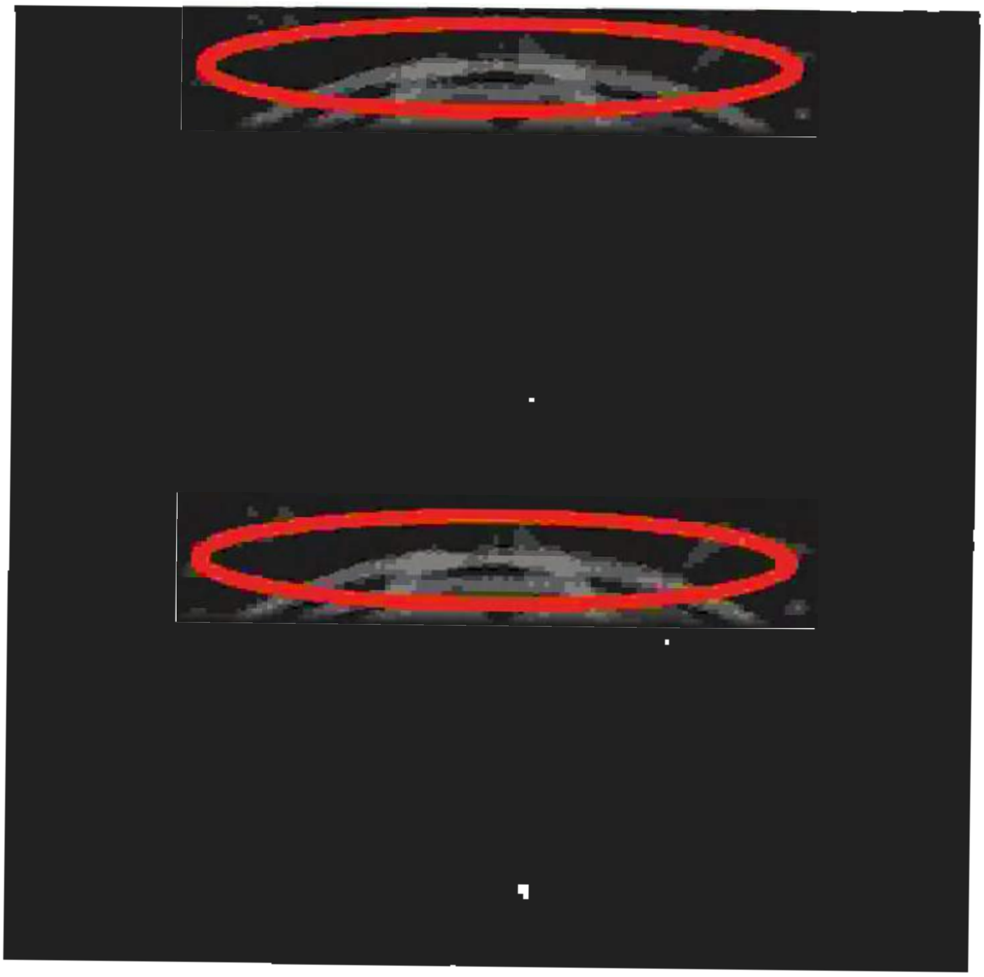


Figure 2.2 *Aliasing effect causes a linear recovery of MRI*

Thus, sparse based algorithms play important role in reconstruction of under-sampled biomedical images. CS theory supports to find the solution of such kind of non-linear recovery of image which is under-sampled. Mathematically, CS recovery algorithms get involves solving under-determined linear equations systems where attained data by scanner (the no. of equations) is less than pixel values of the biomedical image (no. of unknowns).

Sometime given signal or image is sparse in some other domain called sparsifying transform such as wavelet or DCT as described in Figure (2.4). In all such cases, following three main conditions are required for CS theory:

1. **Transform sparsity:** An applied signal or image should have a sparse representation in a certain transform domain.
2. **Incoherence sampling**: In linear reconstruction, the artifacts because of k-space under-sampling must be incoherent in the sparsifying transform domain.
3. **Nonlinear Recovery**: An applied signal or image should be reconstructed through a nonlinear technique that imposes both sparsity of the signal and recovery consistency with the learned samples.

In example of MRI $\mathbf{\Phi} \in \mathbb{R}^{f \times q}$ denotes Fourier matrix and $\mathbf{\Psi} \in \mathbb{R}^{q \times q}$ can be a DCT matrix. Then measurement of CS encoder can be treated as dimensionality reduction problem. From Equation (2.1) and (2.3), measurement vector $\mathbf{y}$ can be expressed as:

$$\mathbf{y} = \mathbf{\Phi}\boldsymbol{x} = \mathbf{\Phi}(\mathbf{\Psi}\boldsymbol{c}) = \boldsymbol{A}\boldsymbol{c} \tag{2.4}$$

Where $\mathbf{A} = \mathbf{\Phi\Psi} \in \mathbb{R}^{f \times q}$ represents a new measurement matrix that maps the signal $\boldsymbol{x}$ into $\mathbf{y}$ measurements in term of sparse representation as shown in Figure (2.3) below:

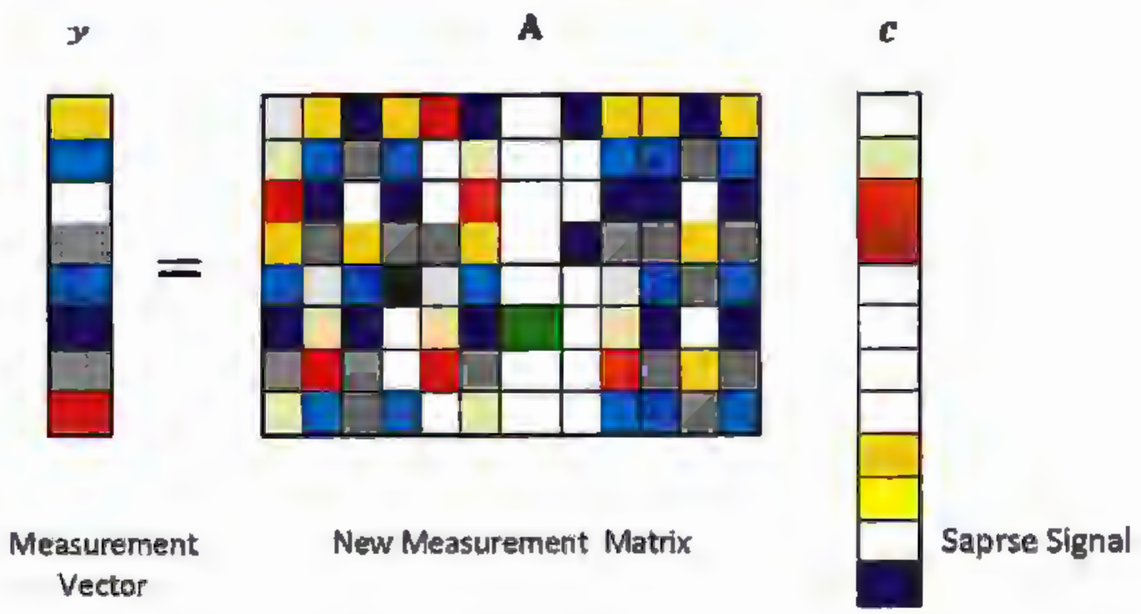


Figure 2.3 *Signal representation in sparsifying domain*

However, for true reconstruction of the signal x, the matrix Φ and reconstruction strategy adopted must satisfy the certain properties. Figure (2.4) explains two different transformations related with medical (MR) image with three domains associated with CS theory.

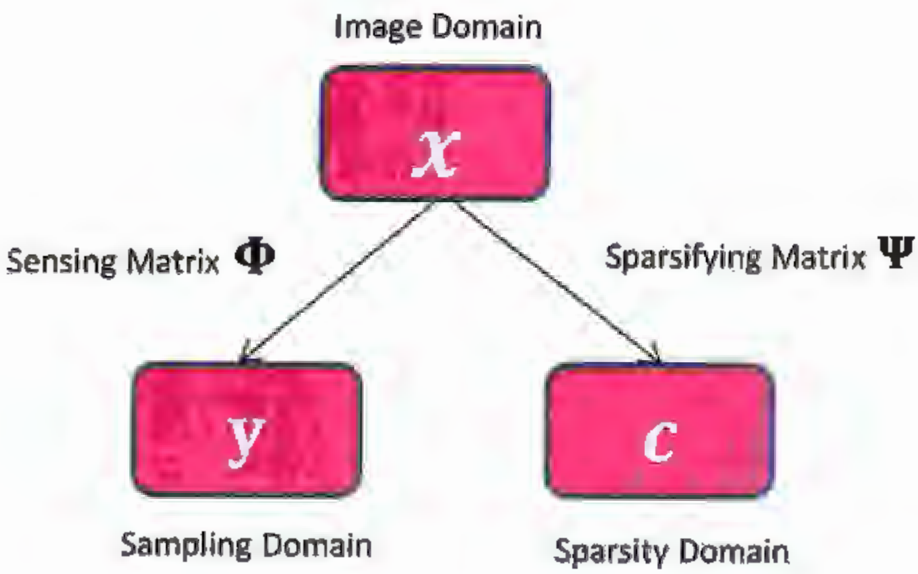


Figure 2.4 *Different domain and transformation in CS*

2.2.1 Sparse Signal Reconstruction

The underdetermined system of linear equations described by Equation (2.4) can have infinitely many solutions indicating that very different signals may lead to the same measurements. So, for a given measurement matrix $\mathbf{\Phi}$ and vector $\mathbf{y}$, finding a maximally sparse solution is an ill-conditioned problem. It is therefore necessary to apply additional regularization constrains to get the desired solution.

The minimum norm solution based on Tykhonov regularization reduces the energy of estimated signal and has a closed form given by [65]:

$$\hat{x} = \min_{x} \|\mathbf{y} - \mathbf{\Phi} x\|_2^2 = \mathbf{\Phi}^{\mathrm{T}} (\mathbf{\Phi}\,\mathbf{\Phi}^{\mathrm{T}})^{-1}\,\mathbf{y} \tag{2.5}$$

However, the energy of the minimum norm solution is distributed over large number of components and is generally non-sparse.

The sparse signal recovery problems such as l_0 minimization use sparsity constraint as regularizer in order to find an estimated solution to Equation (2.3) with limited non-zero entries.

$$\hat{x} = \min_{x} \|\mathbf{y} - \mathbf{\Phi} x\|_2^2 \qquad \mathrm{s.t} \qquad \|x\|_0 \leq \mathrm{K} \tag{2.6}$$

Using the non-convex problem defined by Equation (2.6) to find the sparsest possible solution is computationally impossible [66]. The oldest application of sparse signal approximation is linear regression. Researchers have used several techniques in order to make the sparse signal recovery problems computationally viable.

l_1 norm is convex and has the tendency to sparsify the solution as well. So, a natural approach is to use the convex relaxation and uses l_0 norm in Equation (2.6) by its nearest convex l_1 norm thus recasting the problem as:

$$\hat{x} = \min_{x} \|y - \Phi x\|_2^2 \qquad s.t \qquad \|x\|_1 \leq \varepsilon \tag{2.7}$$

Where ε is a small positive constant value.

Researchers have developed a lot of algorithms to compute the sparse signal approximation to solve the Equation (2.7) e.g. some pursuit methods like BP, BBP[35][67], greedy algorithms like OMP [68], Regularized OMP (ROMP) [69], [70], compressive sampling matching pursuit (CoSaMP) [71] etc. These algorithms provide a faster solution as compared to the convex optimization approach like (LASSO) [72]. There are some Iterative-shrinkage based algorithms (ISA) which got popularity for recovery of sparse signal [73]. Fast iterative shrinkage thresholding (FISTA) [74], iterative-reweighted least squares (IRLS) , separable surrogate functions (SSF) and parallel coordinate descent (PCD) algorithm are few examples of ISA [75].

Other well-known sparse signal reconstruction algorithms include Bayesian sparse learning (BSL) [76], smooth l_0 norm [77] and heuristic algorithms such as message passing [78] , belief propagation [79], particle swarm optimization (PSO) [80].

Solving l_0 minimization problems is NP-hard and conventional PSO has slow convergence to solve sparse signal recovery problem. The Hybrid PSO [81] with sparsity constraints and SSF not only converges to optimal solution quickly but also support of original signal is recovered very precisely. The accuracy in recovery of the sparse signal using proposed hybrid algorithm is far better than that of the signals estimated through PCD and SSF.

2.3 Compressed Sensing MRI (CSMRI)

As per abstract of CS theory, sparse signals and images can be restored under some conditions and constraints [18][61] . MRI is used for medical imaging with a slow data obtaining procedure. The use of CS theory with MRI significantly reduces the scan time along with reduce exposure of ionizing radiation dose to the patient. MRI follows two basic principles for effective CS application:

1) Medical imagery is compressed by sparse coding in a suitable (e.g. wavelet) transformable field.

2) Encoded samples are obtained by MRI scanners, rather than image domain or direct pixel samples.

Protons present in water aids in producing an MRI signal and have a widespread range throughout human body. These photons offer good imaging possibilities. High and good quality images are produced when the MRI signal or k-space frequency data is fully sampled as per criteria of Nyquist-Shannon sampling theorem. However, a great challenge is the acquisition time to acquire these samples. A lot of developments have been made to speed up the long acquisition time over the past few spans of years; however, many technical and physical limits have been reached. One way to reduce the data acquisition time and accelerate the speed in MRI is to get less sampling. But these reduced samples may negatively effect on image quality. A promising theory that can potentially be applied to decrease MRI acquisition time is CS which is considered as one of the most popular techniques for signal compression. CS theory allows under-sampling of k-space below what the Nyquist criteria requires without compromising on image quality. Therefore, CS produces reconstructed images with the same quality in less time.

Cand`es et al. [82] and Lustig et al. [21] introduce and applied CS theory for the good image quality framework and thus improve the imaging speed. So under-sampling for CS must be carefully designed to minimize the errors and artifacts.

More specific to MRI, the CS theory has four potential applications:

(i) Fast angiography

(ii) Improved brain imaging

(iii) Whole- heart coronary

(iv) Dynamic heart imaging.

Parallel imaging (PI) technique is also playing a vital role on hardware side in reduction of acquisition time to combine the signals from multiple coils at the same time permitting some additional parallel sampling of k-space to up lift the image quality [83]–[85]. On the other hand, the final image recovery accuracy depends upon hardware issues like coil receivers calibrations, arrangements/configuration, which is sometimes difficult to attain [20],[86] . Deshmane et al. [87] has discussed and highlighted the benefits of parallel imaging technique for faster acquisition and its effects of under-sampling,

Sparse representation of a signal or medical image is known as highly sparse when it will be represented by minimum information. Usually, MR images are not naturally sparse, however, it can be produced as a sparse representation in transform coefficient domain by using appropriate sparsifying transform and then apply the CS theory for the reconstruction of signal or images.

2.4 Sparsifying Transforms

Methods for CS recovery quality mostly depends on the use of numerous sparsifying transforms such as wavelet, Discrete cosine transform (DCT) or total variation (TV), to recover MR, CT[24] and other biomedical images from the subsampled data by exploiting the sparsity of these images in a transform domain or dictionary. Broadly speaking, these transforms are categorized into predefined or fixed transform and data adaptive transform.

2.4.1 Non-adaptive Sparsifying Transforms

The aim of sparse representation is to estimate a signal by expressing it as a linear combination of some few elementary signals (basis) drawn from a fixed collection (called dictionary) [26],[88]. Speedy computation, a reduced amount of storage and efficient transmission are the main advantages of sparse signal representation. Fixed transform produces better sparse representation, like wavelet, when the given image or signal is piecewise smooth. Otherwise it may not be optimum to use wavelet. Non-adaptive global sparsifying transforms are restricted in conventional MR images to 2–4 folds under-sampling [17],[14],[25],[89]. Numerous unwanted artifacts and loss of features were observed during the reconstruction of images with non-adaptive sparsifying transforms (dictionary) like wavelet and TV, etc. Initially, predefined dictionaries known as non-adaptive sparsifying transforms were used to reconstruct the medical images.

2.4.2 Adaptive Transforms

A lot of work has been conducted to learn adaptive transforms which can better sparsify the images. These transforms are trained to understand the nature of sparsity in an image and acquire an optimally sparse representation regardless of the class of images [27]-[28],[90]. Maingjian Hong et al [91] introduced the SVD based adaptive dictionary learning scheme while Bresler and Ravishanker designed a dictionary learning based MRI [26] . These methods attempts at adaptively make a signal smooth so that transform can be optimally used. Well known adaptive transforms are categorized into synthesis, analytical and sparsifying transform models.

2.4.2.1 Synthesis Transform Model

Adaptive synthesis or analysis dictionaries are broadly used in numerous applications of image processing for example de-noising, in-painting, de-blurring [88],[92]–[97], clustering and classification [98] and tomography [99]. Numerous synthesis dictionaries have been recognized for sparse representation of image such as Curvelet, Contourlets [100] and Ridggelet [101]. In synthesis model similar to linear sparse model, a signal or medical image $\boldsymbol{x} \in \mathbb{R}^{q}$ is represented as linear combination of small numbers of atoms from a dictionary $\mathbb{D} \in \mathbb{R}^{q \times K}$ [102], [103]. Express $\boldsymbol{x} = \mathbb{D}\boldsymbol{\beta}$, where $\boldsymbol{\beta} \in \mathbb{C}^{K}$ is sparse with $\|\boldsymbol{\beta}\|_0 \ll K$, l_0_norm computes the number of non zeros in $\boldsymbol{\beta}$. Practically, adding $\boldsymbol{\varepsilon}$ noise or error in signal domain, above expression can be written as.

$$\boldsymbol{x} = \mathbb{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.8}$$

When $q < K$, dictionary is called over-completed and when $q = K$ and $\mathbb{D}$ is full rank, then dictionary $\mathbb{D}$ is said be a basis. So expression for representing the sparse signal in synthesis model is as follows:

$$\min_{\mathbf{D},\boldsymbol{\beta}} \|\boldsymbol{x} - \mathbf{D}\boldsymbol{\beta}\|_2^2 \ .s.t \ \|\boldsymbol{\beta}\|_0 < \tau_0 \tag{2.9}$$

Where τ_0 is desired level of sparsity .The solution of this synthesis sparse coding is computationally not possible because it is NP-hard (Non-deterministic polynomial-time) and non-convex. The solution may be obtained under some conditions by using some greedy [34],[104], [105] or relaxation algorithm [67], [106], [107]. But those conditions are often violated when synthesis dictionary learning process involves updating its atoms from data itself. To learn a synthesis dictionary from the training signals in Equation

(2.10), well-known method of KSVD and other popular synthesis dictionary learning algorithms faced the convergence problem and stuck in local minima or saddle point during finding solution. Also its sparse coding step is computationally expensive.

2.4.2.2 Analysis Transform Model

In analysis model, some operation or transform domain is desired to make given signal $\boldsymbol{x} \in \mathbb{R}^{q}$ to be sparse. Let operator $\boldsymbol{\Omega} \in \mathbb{R}^{h \times q}$ be the matrix is called analysis dictionary to make $\boldsymbol{\Omega x} \in \mathbb{R}^{h}$ is sparse such that $\|\boldsymbol{\Omega x}\|_0 \ll h$ [108]. Total numbers of zeros in $\boldsymbol{\Omega x}$ is termed as co-sparsity [108]. Finite difference dictionary is considered as well known analysis dictionary. Michal Elad et al. [103] has developed some derivation for the conditions on equivalence of synthesis and analysis based prior.

In practical, analysis model is called noisy signal analysis model when we add noise or error $\boldsymbol{\varepsilon}$ in signal. Then extended version of analysis model can expressed as $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{\varepsilon}$ with $\boldsymbol{\Omega a}$ bieng sparse and $\boldsymbol{\varepsilon}$ is supposed to be very small noise in signal domain [108]. Note that sparse representation $\boldsymbol{\Omega a}$ lies in the range space of $\boldsymbol{\Omega}$. Problem formulation for noisy signal analysis model to recover signal $\mathbf{a}$ will be expressed as follows.

$$\min_{\boldsymbol{a}} \|\boldsymbol{x} - \boldsymbol{a}\|_2^2 \ .s.t \ \|\boldsymbol{\Omega a}\|_0 < \mathrm{h} - \tau_0 \qquad (2.10)$$

Here τ_0 is the minimum allowed co-sparsity. The above Equation (2.4), also called analysis sparse coding problem, is NP-hard and can be solved just like sparse coding in the synthesis model [108],[109]–[113]. This is also computationally expensive. Alternate solutions for Equation (2.10) can also be used to relax the l_0_norm into l_1_norm method with added penalty in the cost function. Both Ophir et al. [114] and Yaghoobi et al.[115] have pursued in analysis dictionary learning to minimize l_0 and l_1 norm of $\boldsymbol{\Omega x}$

respectively. But the problem of convergence behavior is not addressed in those learning techniques. Similarly the other problems, like in synthesis model, such as stuck in bad local minima and computationally expensive persist in analysis model.

2.4.2.3 Sparsify Transform Model

In this model, some operation is desired to make given signal $x \in \mathbb{R}^q$ to be sparse just like analysis model. Let operator $\Psi \in \mathbb{R}^{t \times q}$ be the matrix is called sparsifying transform (or sparsifying dictionary) to approximate $\Psi x \in \mathbb{R}^t$ sparsifiable by using the expression as $\Psi x = \mathbf{x} + \xi$ subject to the condition $\|\mathbf{x}\|_0 << t$, $\mathbf{x}$ is sparse code and ξ is assumed to be very small residual (noise or error) in transform domain. A classical alternate representation of transform sparsity is described through following Figure (2.5).

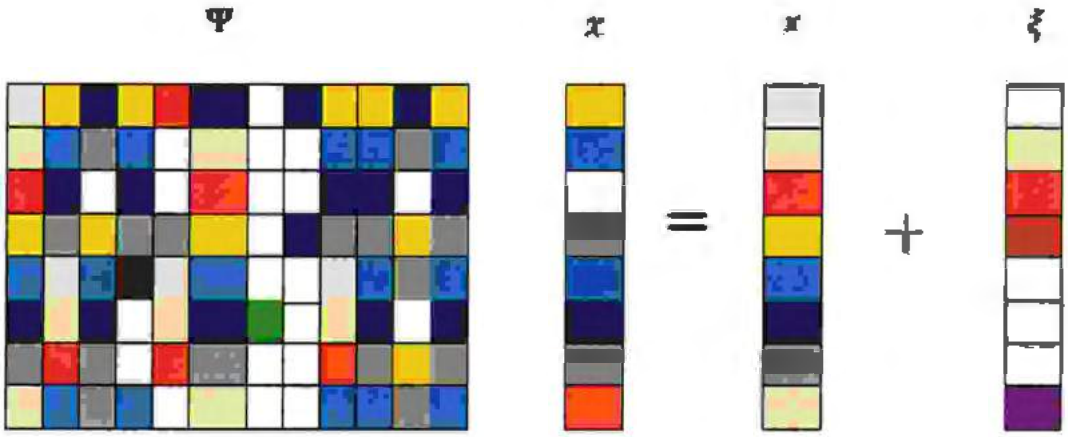


Figure 2.5 *Transform Sparsity of a signal*

Here two important points lead the difference between sparsifying transform and last discussed two models (i.e. synthesis and noisy signal analysis models).

1) The error or noise ξ is in transform domain in sparsifying transform whereas the error ε is studied on lying signal domain in synthesis or analysis model.

2) Unlike synthesis and analysis model, the sparse representation of x is not forced to lie in range space of Ψ(sparsifying transform)

So we can formulate the noisy sparsifying transform problem as follows.

$$\Psi x = \Psi(a + \xi) = \Psi a + \Psi \xi = x' + e' \quad (2.11)$$

Where $\Psi\xi = e'$ is transform domain noise or error of Ψ and $\Psi a = x'$ represents transform code with error e' for the noisy x. Although, if a signal x satisfy the noisy signal analysis then it will be true to an equivalent transform model but the contrary is not correct in general. Because spare signal code is lie in range space of Ω. We may conclude that (noisy signal) analysis model is typically more constrained than transform model. For known sparsifying transform ψ, and given signal x , then we express transform sparse coding problem as

$$\min_{x} \|\psi x - x\|_2^2 \qquad s.t \ \|x\|_0 < \tau_0 \quad (2.12)$$

Signal can be recovered almost accurate by thresholding the product Ψx and retaining the τ_0 components of leading magnitude, while setting the rest of components to zero. Recovered signal can also be obtained from least squares estimate for given transform and sparse code. Therefore unlike synthesis and analysis model, transform model permits to exact and fast computation along with guaranteed convergence.

2.5 Dictionary Learning

A dictionary represents a matrix of values which when applied on a MRI data can produce a sparse form of the image or signal. Sparsifying transforms have an analytical form of dictionary (like Wavelet, DCT and FT) and are fast in execution and implementation but they generally lack behind in sparsifying signals. Such transform are appropriate only for a class of signals at best and manage sub optimal performance for others. So underlying cause get involve into basis functions simply. Therefore, there is a requirement to implement "adaptability" of dictionary learning techniques which can provide improved results in sparsifying an image or signal.

Dictionary can be learned into following two ways:

- Whole image dictionary learning
- Patch based dictionary learning

2.5.1 Whole Image Dictionary Learning

In this framework, dictionary learning and sparse coding are alternatively updated for whole image as shown in Figure (2.6).

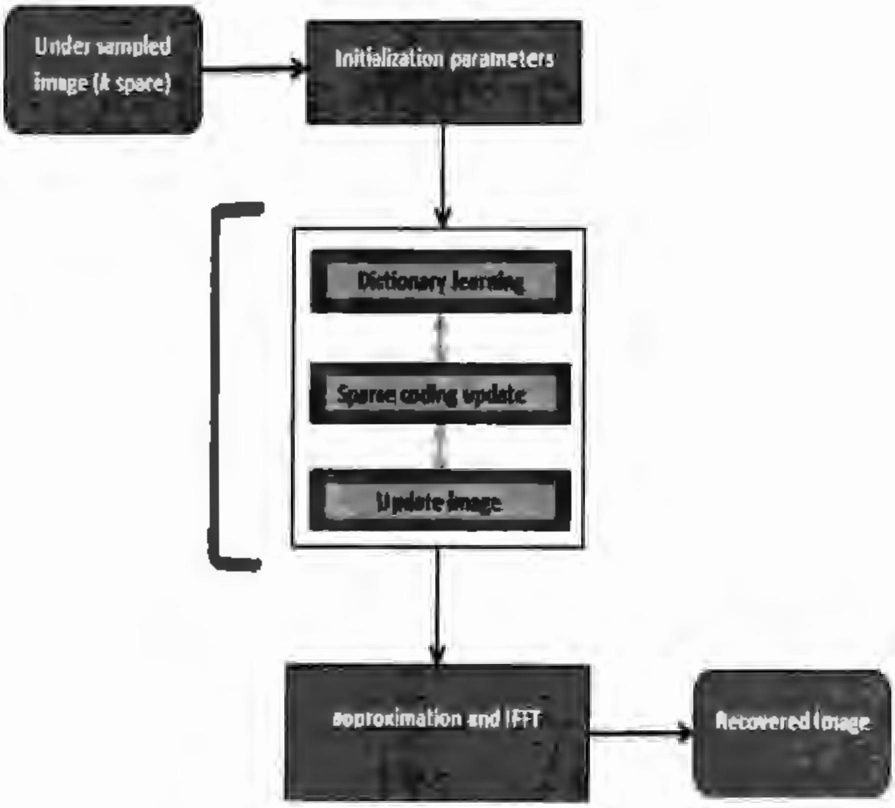


***Figure 2.6** Dictionary Learning framework for whole image*

2.5.2 Patch Based Dictionary Learning

It has been established by recent research work, dictionary learning from small patches of an image gives superior results than form a reference (whole) image[26].

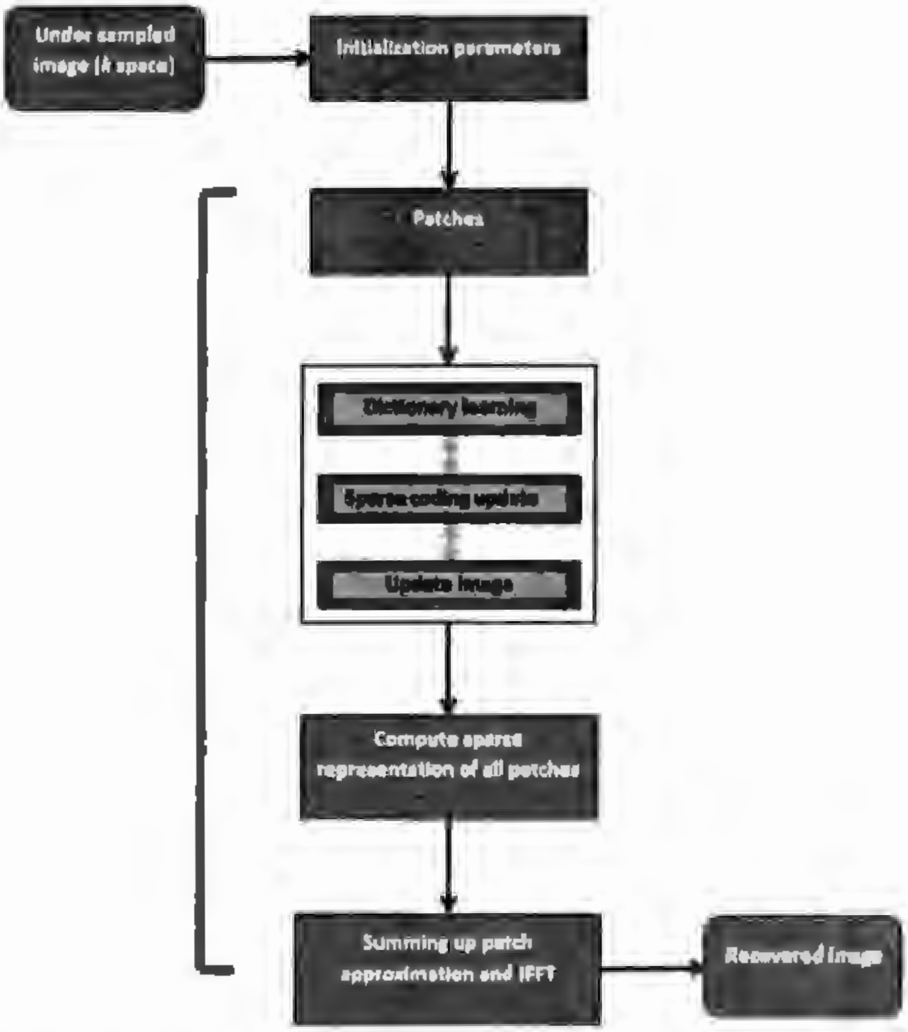


Figure 2.7 *Patch based dictionary learning framework for brain MR image reconstruction based on CS.*

In patch based dictionary learning method, given image is dissolved into small image patches and then learn the dictionary for each patch. Then sum all the patches after learning the dictionary for each patch, so the combined reconstruction and dictionary updates problem is solved iteratively, where each iteration is involved two steps

1. Sparse recovery
2. Update dictionary atoms

Figure (2.7) describes the detail steps during the reconstruction of MR image using patch based dictionary learning.

2.6 Representation of Under-sampled MR Image

The aim of the sparse approximation is to synthesize a given signal or measurement vector as a linear combination of a small number of sparsifying transform vectors , ψ_i.

MR image acquisition can be modelled as an under-sampled measurement of MR image in k-space with the help of measurement matrix $\mathbf{\Phi}_u$. Mathematically, MR image $\boldsymbol{x} \in \mathbb{C}^q$ is encoded to a measurement vector $\boldsymbol{z} \in \mathbb{C}^m$ as

$$\boldsymbol{z} = \mathbf{\Phi}_u \boldsymbol{x} \tag{2.13}$$

Where $\mathbf{\Phi}_u \in \mathbb{C}^{m \times q}$ is under-sampled Fourier encoding or measurement matrix. Whenever the number of unknowns is greater than the number of k-space samples $(q > m)$, it is called under-sampling.

Compressed sensing (CS) provides a promising way of reconstructing $\boldsymbol{x}$ from its undersampled measurements $\boldsymbol{z}$ provided $\boldsymbol{x}$ is sparse in some sparsifying transform domain, $\mathbf{\Psi}$, also known as the dictionary. The recovery of $\boldsymbol{x}$ can be formulated as l_0 minimization of sparsified image $\mathbf{\Psi}\boldsymbol{x}$ [20]:

$$\min_{\boldsymbol{x}} \|\mathbf{\Psi}\boldsymbol{x}\|_0 \qquad \text{s.t.} \quad \mathbf{\Phi}_u \boldsymbol{x} = \boldsymbol{z} \tag{2.14}$$

$\Psi \in \mathbb{C}^{t \times q}$ is a sparsifying transform for the image such as wavelet or DCT or any other learned dictionary.

The disadvantage of the model in Equation (2.14) is that the sparse coding step is NP (Non-deterministic Polynomial-time) hard because the algorithm involves $l_0 -$ quasi norm. However, solution can be achieved through some greedy algorithm like OMP [68], [116], linear programming [117],basis pursuit de-noising [118] and some other empirical based algorithms [119]–[122] for such kind of sparse coding problem. The non-convex formulation of Equation (2.14) can be transformed to a convex problem by using $l_1 -$ norm ,i.e;

$$\min_{x} \|\Phi_u x - z\|_2^2 + \gamma \, \|\Psi x\|_1 \tag{2.15}$$

Where γ is the Lagrangian multiplier.

CS has been successfully applied in the recovery of static and dynamic MR images [23],[123],[124],[125] but in this thesis we focus on static MR images and study it in detail.

Saiprasad et al. [26] has presented his the new theory over CSMRI based on dictionary learning , known as dictionary learning based MRI (DLMRI). DLMRI is one of the leading techniques to recover the MR images by learning sparsifying transform.

2.7 Dictionary Learning based MRI (DLMRI)

In CSMRI, the theory CS exploits the sparsity of medical signal or signal patches in a transform domain or synthesis dictionary to reconstruct signals from few or compressively measurements. But, DLMRI method reconstructs the basic signal(s) simultaneously as well as learns the sparsifying transform from highly compressively measurements.

Saiprasad et al [26] explains DLMRI scheme based on K-SVD for learning sparsifying transform with OMP for sparse coding. The K-SVD updates the columns sequentially one by one and takes time to update all the atoms. This technique also faces the increase of singular vectors. Although the K-SVD performs well in capturing a reference dictionary, but its denoising performance is comparatively limited. The artifacts at the edges appear due to the hard constraint applied by OMP. The hard sparseness constraint may not be relatively capable for medical imaging application which produces artifacts on high frequency. There are also some other methods, such as method of optimal directions (MOD), iterative least square (ILS), recursive least square (RLS) and simultaneous coded optimization (SimCO), alternating direction method of multipliers (ADMM) for dictionary learning and basic pursuit (BP), least absolute shrinkage and selection operator (LASSO), and focal underdetermined system solver FOCUSS for sparse coding.

2.8 Block Dictionary Learning

One of the key characteristic of MRI is that MR signal can be extracted from union of a small number of subspaces. Dictionary atoms are categorized into underlying subspaces in such kind of signals which lead to sparse representation for block sparse structure [20]. There are many methods, such as block basic pursuit (BBP, block orthogonal matching pursuit (BOMP), and group Lasso that have been suggested to get the benefit of this structure in recovering the block sparse representation. Predetermined dictionaries and known block structures are supposed to be used in these methods. Figure (2.8) shows the two equivalent models for block dictionaries $\mathbb{D}$ with block structures b having 5 blocks, together with 2-block-sparse representation of $\boldsymbol{\beta}$. Both models represent the same signal, since the atoms in $\mathbb{D}$ and values of **b** and $\boldsymbol{\beta}$ are permutated in same manners. In this example, atoms of dictionary $\mathbb{D}$ are sorted and organized into blocks, and which delivers an accurate set of representation vectors whose non-zeros entries are focused to a fixed amount of blocks.

Block structured dictionaries are learned to utilize any embedded structure in order to produce an effective sparse approximation. Method for block-structured dictionary involves two steps: update the dictionary block and then finding the sparse coefficients according to block structure in the dictionary. Algorithm BOMP picks the blocks from dictionary sequentially best to input applied signals. BLKSVD technique updates the atoms in blocks and corresponding non-zero coefficients simultaneously. Surprisingly, learning of block transforms or block dictionary has received no attention for recovery of medical images e.g. for MRI. Figure (2.8) shows two equivalent models of block dictionaries $\mathbb{D}$ and block clusters **b** having 5 blocks. Both models represents same signal.

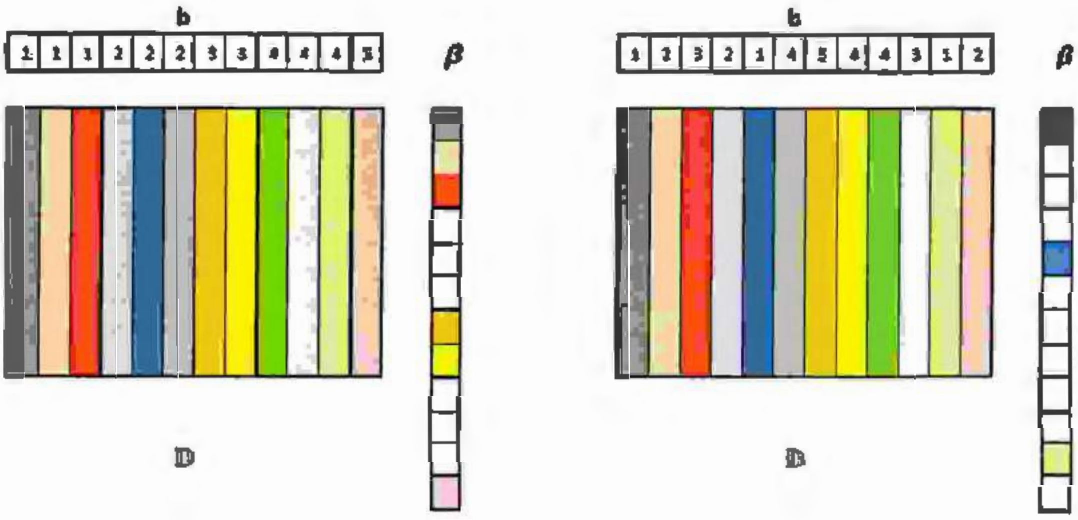


Figure 2.8 *Two equivalent models of block dictionaries*

2.9 Dictionary Learning Challenges:

Some main challenges for dictionary learning are as follows:

(1) Training samples are imperfect: The training samples are noisy and distorted during sensing process. It may be partial or incomplete such as case in multi-view imaging. So it becomes more challengeable for dictionary learning when the underlying causes of signals or the related data in observations is noisy, distorted or missing.

(2) Nonlinear acquisition techniques and method: Linear models come to be invalid in advanced applications, for example in, medical imaging where the acquisition techniques are normally nonlinear.

(3) Application based selection of proper sparse representation method: The characteristics of the signals and the processing task play a vital role in learning a good dictionary. A dictionary can be good for representing a class of signals, but choosing the right sparsity techniques or sparse coding technique is playing the key role in reconstruction of images typically medical imaging.

General speaking, dictionaries offer a very flexible and powerful approach to symbolize related data and information in high-dimensional signals. However, the appropriate and suitable modeling of the complex underlying causes of observations holds many sensational and exciting questions regarding the appropriate construction of these dictionaries.

2.10 Performance assessment parameters

Recovered MR images are assessed through the quality and performance parameters like PSNR, high frequency error norm (HFEN), correlation, SSIM (Structural SIMilarity) and sharpness index (SI).

PSNR is calculated, normally defined in decibels (dB), as the ratio of the maximum possible intensity level of the original image to the root mean square (RMS) compressed/reconstruction error relative to the original image. On image compression, it is considered as a standard image quality measure and is being used in compressed sensing MRI before [126], along with the associated metric of the signal to noise ratio (SNR in dB).

Reconstruction of edges quality and fine features are measured by (HFEN) which is calculated as the norm of the result acquired by Laplacian of Gaussian (LoG) filtering, the difference between the reconstructed and reference images.

SSIM measure the similarity between two MR images, original x and estimated $\hat{x}$. When maximum value of SSIM is approximately equal to 1, this indicates that both images are almost similar and it is calculated using following relationship.

$$SSIM(x,\hat{x}) = \frac{(2\Omega_x\Omega_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\Omega_x^2 + \Omega_{\hat{x}}^2 + C_1)(\sigma_x^2 + x_{\hat{x}}^2 + C_2)} \tag{2.16}$$

Where $C_1$and C_2 represents the constants which depends on the dynamic range of the images. Ω_x and $\Omega_{\hat{x}}$ shows the mean values whereas σ_x^2 and $\sigma_{\hat{x}}^2$denote the variances of the given (original) and reconstructed image correspondingly. $\sigma_{x\hat{x}}$ is the covariance between of given and recovered image.

Sharpness index (SI), a measuring parameter based on image sharpness is given as [127].

$$SI(x) = -log\phi\left(\frac{m - TV(x)}{v}\right) \tag{2.17}$$

Where $\mathbb{m} = \mathbb{E}[(TV(x))]$ is the expected value of the total variation of the recovered image $\hat{x}$, $v = \mathbb{V}[(TV(x))]$ is the corresponding variance, and φ represents the normal distribution [127].

We have performed our simulation on Matlab 9.2.0.538062 (R2017a). Computations were performed with 7th Generation Intel Core i5-7200U Processor (2 Cores - 4 Threads).

2.11 Summary

In this chapter, we have discussed the compression technique for the recovery of MR images based on CS theory with its importance. Learned sparsity models are valuable and well-founded tools for modeling data due to computational and performance advantages. Therefore learning sparsifying transform used in imaging leads to state of the art results. Classification and different models of dictionary learning are discussed to represent a MR image as a linear combination form and its sparse representation. Finally, different performance parameters are discussed to evaluate the quality of recovered images.

Chapter 3

MR Images Reconstruction Using Dictionary Learning Technique

Compressed sensing (CS) is a fascinating and rapidly growing arena in numerous applications of image processing such as biomedical imaging, as it allows an accurate reconstruction of images through a rational, by exploiting the image sparsity and mainly depends on the use of diverse sparsifying transforms like wavelets, curvelets or TV to recover MR images. According to a developed mathematical concept of CS, the biomedical images with sparse representation can be reconstructed from randomly under-sampled data. Because of high under-sampling, the reconstructed images are suffered due to noise like artifacts. Reconstruction of images from CS involves two steps, one for dictionary learning and other for sparse coding. In this chapter we discuss such a framework, where we choose SimCO patch-based dictionary learning that updates the atoms simultaneously whereas FOCUSS is used for sparse representation because of soft constraint on sparseness of an image. Merging SimCO and FOCUSS, we come up with a new scheme called SiFo. Our suggested reconstruction scheme learns the dictionary (uses it to eliminate aliasing and noise in one stage) and afterwards restores and fills in the k-space data. Series of experiments were carried out, using different sampling schemes, on noisy and noiseless cases of phantom as well as real brain images. Based on various performance parameters, it is evident that our designed technique overtakes the conventional techniques like K-SVD with OMP, used in dictionary learning based MRI (DLMRI) reconstruction.

3.1 Patch Based Representation of MR Image

Reconstructed image quality mainly relies on the sparsifying transform. The main constraint of high under-sampling in non-adaptive compressed sensing of MR images is solved by adaptive dictionary updates with high sparsity. In our framework, we use patch-based adaptive dictionary learning. For this purpose, let the given image $\boldsymbol{x} \in \mathbb{C}^{q}$ be denoted as a combination of patch vectors $\boldsymbol{x}_{ij} \in \mathbb{C}^{n}$ of 2D squared image where dimension of each patch is $(\sqrt{\mathrm{n}} \times \sqrt{\mathrm{n}})$ pixels. (i, j) marks the position of a patch starting from top left corner of the image. $\mathbf{D} \in \mathbb{C}^{\mathrm{n} \times \mathrm{K}}$ denotes the patch-based dictionary having K number of atoms and $\boldsymbol{\theta}_{ij} \in \mathbb{C}^{\mathrm{K}}$ is the sparse representation of $\boldsymbol{x}_{ij}$ patch with respect to $\mathbf{D}$. Dictionary $\mathbf{D}$ is said to be over complete when $\mathrm{n} < K$.

The following optimization problem solves the patch-based dictionary learning as

$$\min_{\mathbf{D},\mathcal{G}} \sum_{ij} \left\| \mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij} \right\|_2^2 \quad \text{s.t.} \left\| \boldsymbol{\theta}_{ij} \right\|_0 \leq \tau_0, \quad \forall\, i, j \tag{3.1}$$

where $\mathcal{G}$ represents the set $\{\boldsymbol{\theta}_{ij}\}_{ij}$ of sparse approximation of all patches and τ_0 is the required sparsity. $\mathbf{W}_{ij} \in \mathbb{R}^{\mathrm{n} \times \mathrm{q}}$ matrix acts as an operator that brings out the patch $\boldsymbol{x}_{ij}$ from given image $\boldsymbol{x}$ such as

$$\boldsymbol{x}_{ij} = \mathbf{W}_{ij}\boldsymbol{x} \tag{3.2}$$

This learning schemes in Equation (3.1) helps to minimize the total fitting error of all image patches while learning the dictionary, subject to sparsity restrictions.

Optimization formulation used in Equation (3.1) is NP hard for the fixed $\mathbf{D}$ and can be solved from many algorithms like MOD, K-SVD[26] or SimCO [27],[26],[33] . Such kind of algorithms normally alternate between finding the dictionary **D**, and sparse representations $\mathcal{G}$. Most researchers use the K-SVD algorithm, in which the dictionary atoms are updated sequentially along with related sparse coefficients for the patches. Since singular value decomposition (SVD) is used k-times in this algorithm for updating the k-atoms of the dictionary sequentially, hence called K-SVD.

3.2 Problem Formulation

Reconstructed compressively sampled biomedical images typically suffer from numerous artifacts on high under sampling factors. Under-sampling of k-space and noise in samples are two main causes of artifacts. A decent dictionary must be capable of minimizing the artifacts which are noticed in zero filled Fourier reconstruction and be consistent to produce reconstructed images by available k-space data. Possible cost function is as follows:

$$\min_{\mathbf{D},\mathcal{G},\boldsymbol{x}} \sum_{ij} \left\| \boldsymbol{x}_{ij} - \mathbf{D}\boldsymbol{\theta}_{ij} \right\|_2^2 + \eta \left\| \boldsymbol{\Phi}_u \boldsymbol{x} - \boldsymbol{z} \right\|_2^2 \qquad \text{s.t.} \left\| \boldsymbol{\theta}_{ij} \right\|_0 \leq \tau_0, \ \forall i,j \qquad (3.3)$$

In Equation (3.3), the 1st term is responsible for the quality of the sparse approximation of the patched images with respect to the dictionary **D** whereas 2nd term enforces data consistency in k-space. Parameter η depends on standard deviation σ of measurement noise such as $\eta = (\frac{\lambda}{\sigma})$ and λ is taken as a positive constant.

Proposed cost function is capable of learning an adaptive dictionary to reconstruct the underlying image but it is NP hard and non-convex even when l_0 − quasi norm is relaxed to l_1 − norm. Hence we use alternate minimization methods to solve this problem. Elimination of aliasing and noise is done through the adaptive patched sparsity whereas the elimination of the artifact is done from overlapping patches.

3.3 The SiFo Algorithm

Adaptive learning of sparsifying transforms is a highly non-convex problem and is computationally expensive. Typically the problem in Equation (3.3) is solved in two steps. (i) Dictionary learning and sparse coding are updated alternately keeping the estimated signal x fixed while in second step (ii) the estimated signal x is updated to satisfy the data fidelity while keeping dictionary and sparse representation fixed.

3.3.1 Updating the dictionary and sparse coding

Since x is fixed, the objective function in Equation (3.3) becomes

$$\min_{D,\mathcal{G}} \sum_{ij} \left\| x_{ij} - D\theta_{ij} \right\|_2^2 \quad \text{s.t.} \|d_k\|_2 = 1 \ \forall k \qquad \& \ \left\|\theta_{ij}\right\|_0 \le \tau_0, \ \forall i,j \qquad (3.4)$$

Extra constraint of unity norm on dictionary atoms (d_k, $1 \le k \le K$) is applied to avoid scaling issues [123]. Many researchers use the K-SVD to learn the dictionary where the atoms of the dictionary are updated one by one i.e. K-times SVD which increases the computation time. Saiprasad et al [26] performed tremendous work on MR image reconstruction from highly under-sampled k-space data using the K-SVD technique to learn the dictionary and a greedy algorithm such as OMP for updating sparse coefficients. His working on DLMRI showed noticeable improvements in the reconstruction of different medical images along with other performance parameters such as SNR and high frequency error numbers (HFEN). He compared his results with Lustig et al [21] (denoted as LDP). One main problem in OMP is that it imposes the hard constraint [46] to achieve sparse solution. This hard sparseness limitation may not be suitable for medical imaging application because fast and sudden changes of the medical signal values typically on high frequency artifacts. Other problem for OMP algorithm is that it is not very computationally efficient.

We have used the FOCUSS for updating the sparse representation matrix. Since the FOCUSS introduces the sparseness of the image as a soft constraint, high frequency artifacts are also minimized as compared to OMP because the non-zero image values are progressively suppressed. FOCUSS also inclines to repress the recovery noise as aliasing artifacts and noises are typically isolated; thus, these can be simply detached through the pruning process of this algorithm. The empirical results show better outcomes especially in noisy case. Lastly, FOCUSS can be applied computationally in an efficient manner by means of successive quadratic optimization. This is a relatively significant advantage over computationally expensive sparse optimization algorithm such as OMP.

In this framework, we use SimCO [33] to learn the dictionary **D**. The main attribute of SimCO is to update all the atoms and corresponding non-zero coefficients simultaneously and hence reduce the computation cost.

Because of above mentioned advantages of SimCO and FOCUSS for reconstruction of MR images, we term our framework "SimCO plus FOCUSS" the SiFo. Experimental results on synthetic as well as in vivo data determine improved performance of our algorithm even from highly sparse k-space samples and show better results than Saiprasad et al [26].

3.3.2 Updating the estimated image(s) for reconstruction

To update the reconstruction image $\boldsymbol{x}$, keep the dictionary and the sparse representation constant then the sub-problem for our cost function in (3.3) can also be written as follows:

$$\min_{\boldsymbol{x}} \sum_{ij} \left\| \mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij} \right\|_2^2 + \eta \left\| \boldsymbol{\Phi}_{\mathrm{u}}\boldsymbol{x} - \boldsymbol{z} \right\|_2^2 \quad (3.5)$$

The formulation in Equation (3.4) is the least square problem and detail solution is in appendix A. The solution is as follows.

$$\boldsymbol{\Phi x}(k_x, k_y) = \begin{cases} \mathcal{N}(k_x, k_y) & ,(k_x, k_y) \notin \mho \\ \dfrac{\mathcal{N}(k_x, k_y) + \eta \mathcal{N}_0(k_x, k_y)}{1+\eta} & ,(k_x, k_y) \in \mho \end{cases} \tag{3.6}$$

Here $\boldsymbol{x}$, is reconstructed by taking the IFFT of $\boldsymbol{\Phi x}$. From (ix) in appendix A

$$\mathcal{N} = \boldsymbol{\Phi} \sum_{ij} \mathbf{W}_{ij}^{T} \mathbf{D}\, \theta_{ij} \frac{1}{\alpha} \tag{3.7}$$

Equation (3.7) is called the "patched average result" in Fourier domain and $\boldsymbol{\Phi x}(k_x, k_y)$ denotes the restructured updated value on position(k_x, k_y) of the k-space. $\mathcal{N}_0 = \boldsymbol{\Phi}\ \boldsymbol{\Phi}_u^H \boldsymbol{z}$ expresses zero- filled k-space measurement and $\mho$ denotes the subset of k-space that has been sampled.

Algorithm 3.1:

Goal: To learn the dictionary for reconstruction of under-sampled image

Input: z= training signal in k-space measurements, μ, p

Output: $\hat{\boldsymbol{x}}$ = An estimated reconstruction MR Image

Initialization: $\boldsymbol{x}=\boldsymbol{x}_0 = \boldsymbol{\Phi}_u^H \boldsymbol{z}$

Main Iteration:

1. Alternately learn dictionary by SimCO and sparse (coding) approximation for $\boldsymbol{x}$ patches by FOCUSS.
2. Update $\boldsymbol{x}$
3. $\mathcal{N} \leftarrow \mathcal{FFT}(\boldsymbol{x})$
4. Restore sampled frequency to update the $\mathcal{N}$ as per (3.5)
5. $\hat{\boldsymbol{x}} \leftarrow \mathcal{IFFT}(\mathcal{N})$

More extensive pseudocode is presented in the appendix-B1

The complete process for the proposed technique SiFo is described as below in the Figure (3.1).

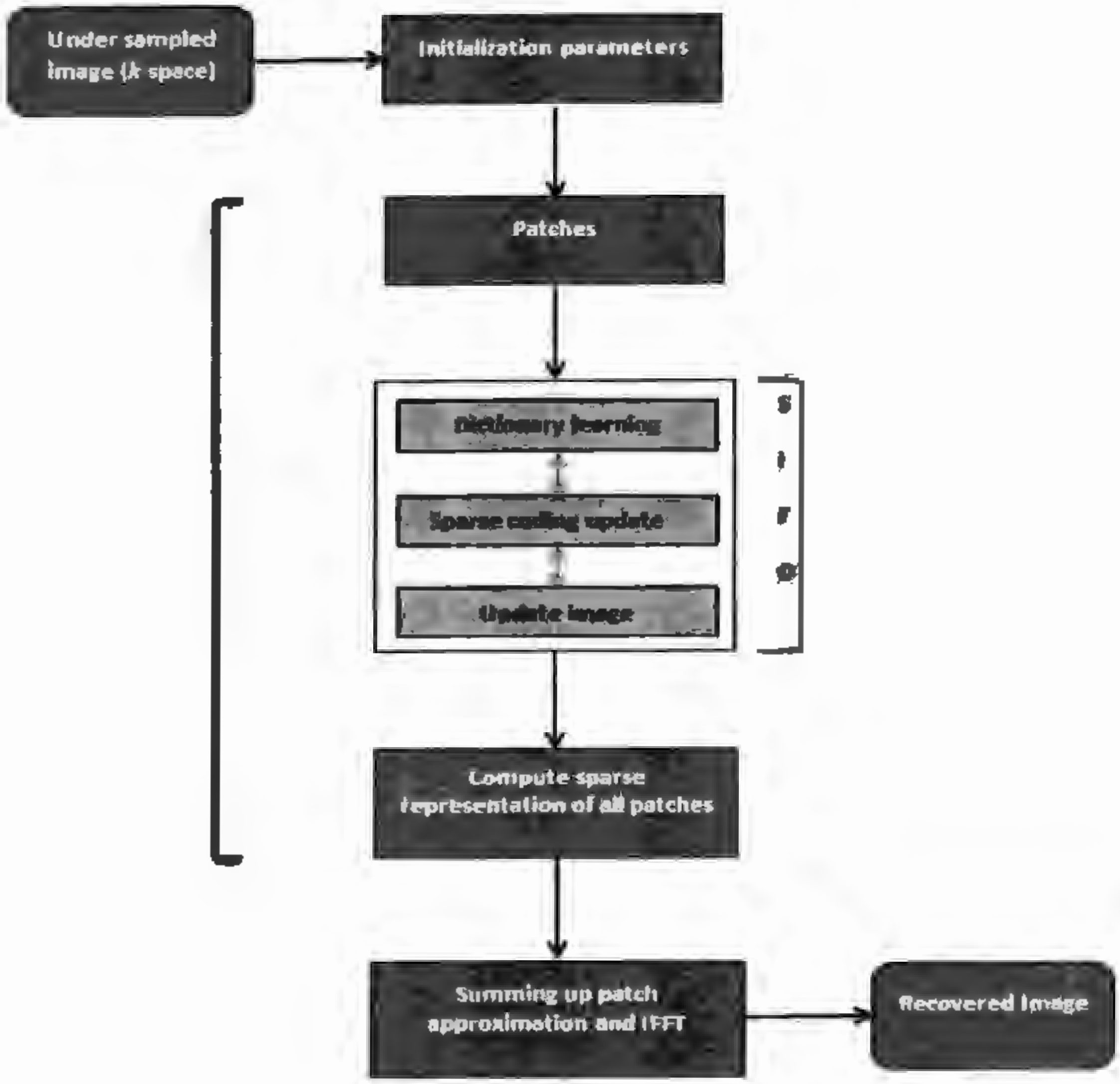


Figure 3.1 *Patch based dictionary learning framework for brain MR image reconstruction based on CS through SiFo method*

3.4 Empirical Results and Discussion

One of the important factors in dictionary learning is its initialization. In these experiments, we initialize our dictionary from subset of image patches. In our experiment we take the image in Fourier domain, and apply our proposed framework for noiseless and noisy scenarios. Image in Fourier data is processed with different sampling schemes. These images are converted to over lapping patches of size $(\sqrt{n} \times \sqrt{n})$, in our case n = 36. From these patches, we have initialized the data for dictionary learning.

The dictionary is updated through regularized SimCO and sparse coefficients using FOCUSS. Through different iterations, our proposed algorithm reconstructs the image properly which outperform the DLMRI proposed by Saiprasad et al. Results are shown in Figures (3.2-3.7) and Tables (3.1-3.3). The DLMRI technique based on KSVD has outperformed several methods like MOD and LDP (method by Lusting et al). Therefore in our comparison, we compared our method with only DLMRI. Since our proposed method performs better than the DLMRI, hence it will also perform better than other CSMRI technique. So our method SiFo for reconstruction is compared with a leading DLMRI method by Saiprasad et al [26].

The performance of the proposed algorithm is validated with various under sampling factors, for noiseless and noisy cases. The under-sampling is directly applied on k-space (fully-sampled) MR data set. Axial T2-weighted reference images of the brain are used as MR images, taken from vivo MR scans of size 512×512 from American Radiology Services (ARS).

During the tests with noisy and noiseless images, we fixed values of different parameters such as atoms $K = n = 36$, sparsity $\tau_0 = 6$, $\gamma = 140$, maximum overlapped patch called overlap stride "r" as $r = 1$ (the distance in pixels between the corresponding pixel position in adjacent image patches.), regularized parameter for SimCO is $\mu = 0.05$ and for FOCUSS diversity measure $p = 0.5$. Both SiFo and DLMRI are performed for 15 iterations with fixed sparsity τ_0 and 200xK patches.

The SiFo and DLMRI learning techniques need an initialization as discussed above for the dictionary. Real valued sparsifying transforms [128] were used in the simulated experiments with real valued images. All performance tests are employed on Shepp-Logan Phantom as well as real brain MRI data.

3.4.1 Performance in Noiseless Scenario

We first compared the proposed method with the DLMRI in noiseless case. Figure (3.2) shows that the proposed algorithm has shown better results in the noiseless scenario. Algorithm performance on a phantom and a brain image is evaluated using 2D variable density random sampling of k-space. The dictionary learning algorithm (SiFo) reconstructed both the images free from artifacts and aliasing effect. The results were achieved by running our algorithm for 15 iterations. The reconstruction with SiFo algorithm is clearer and sharper than with DLMRI.

3.4.1.1 PSNR

From the Figure (3.2) and Table (3.1), the comparison of PSNR for both methods can be analyzed which shows that SiFo converges quickly by using $l_2 -$ norm reconstruction of difference between two successive iteration. So our algorithm is appropriately minimizing the noise and aliasing observed in the zero filled case to delivers a better reconstruction.

3.4.1.2 HFEN

The SiFo performs better to capture the image of brain and phantom with fast convergence than DLMRI. This can be easily observed from Figure (3.2) and (3.3), where proposed method showed better edges and sufficient features of reconstruction images.

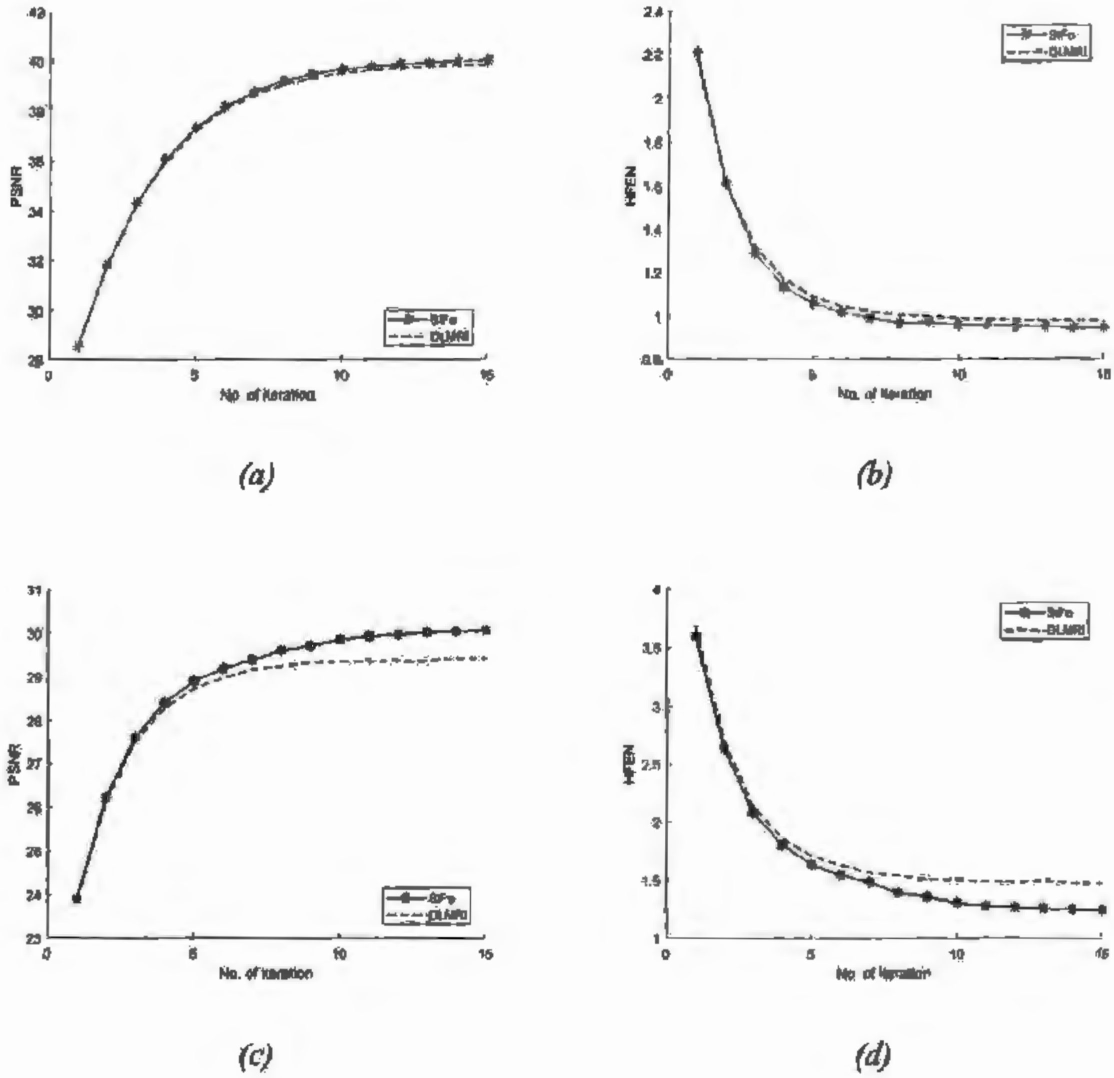


Figure 3.2. *Algorithm performance in noiseless case. (a) PSNR vs iterations with comparison to DLMRI for brain image (b) HFEN vs iterations with comparison to DLMRI for brain image (c) PSNR vs iterations with comparison to DLMRI for phantom image (d) HFEN vs iterations with comparison to DLMRI for phantom image.*

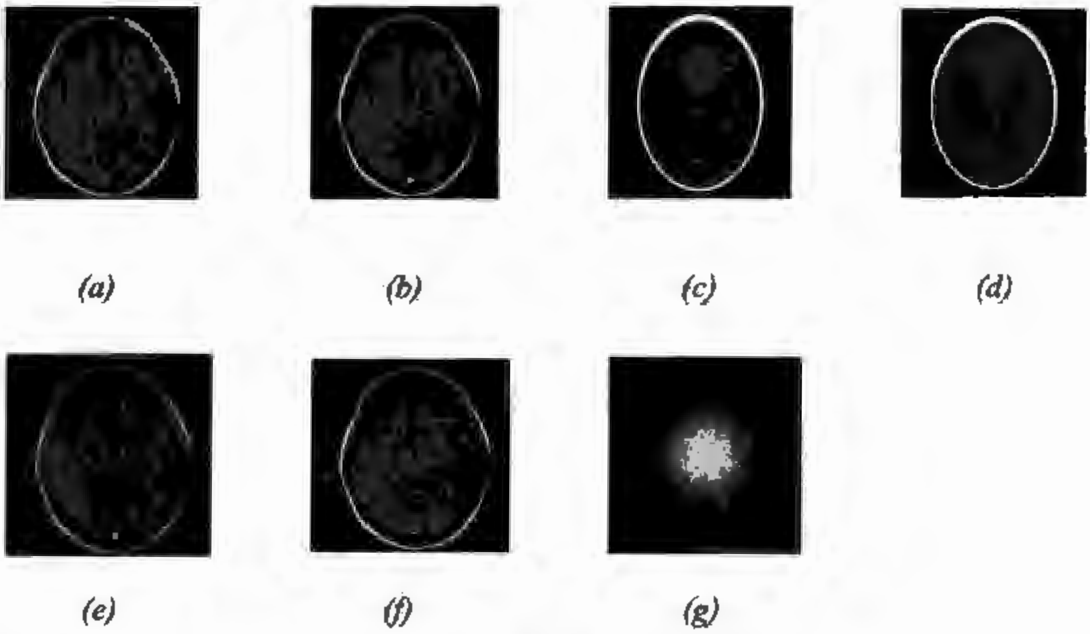


Figure 3.3. *Images recovery for noiseless case. (a) Recovered MR image of brain by SiFo (b) Recovered MR image of brain by DLMRI (c) Recovered MR image of phantom by SiFo (d) Recovered MR image of phantom by DLMRI (e) Reconstruction brain image with zero filling (f) Reference MR image for brain (g) k-space sampling mask with 10 fold.*

3.4.1.3 Correlation/Similarity index/Sharpness:

Although the correlation and similarity index display marginal improvement in the reconstruction of images for brain and phantom in noiseless case, the sharpness indicates good results as shown in Table (3.1).

(a) Noiseless case for brain Image					
No	Parameters	DLMRI	SiFo	Difference	Remarks
1	**Correlation**	0.998	0.9881	0.0001	improved
2	**Similarity Index (SSIM)**	0.8899	0.8935	0.0036	improved
3	**Sharpness**	944.1884	984.3163	40.1279	Improved
(b) Noiseless case for phantom Image					
No	Parameters	DLMRI	SiFo	Difference	Remarks
1	**Correlation**	0.988	0.9896	0.0016	improved
2	**Similarity Index (SSIM)**	0.8151	0.8492	0.0341	improved
3	**Sharpness**	3739	4434.4	695.4	improved

Table 3.1. *Performance parameter (SiFo vs DLMRI) of algorithm with noiseless case for (a) Brain image (b) Phantom image.*

3.4.2 Performance in Noisy Scenario

In this case, we added zero-mean white Gaussian fixed noise of standard deviation for all cases such that sigma =10.2489 in k-space data. In the course of reconstruction update stage of the algorithm in Equation (3.5), the noisy scenario involves weighted averaging in k-space. The performance of our technique on the fully sampled noisy image is observed by using different sampling mask from Figures (3.4-3.7) and Tables (3.2-3.3). Our process of reconstruction has sufficiently removed the noise and aliasing observed in zero-filled result.

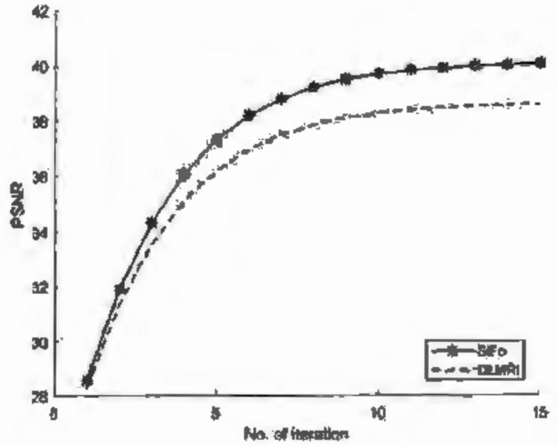
3.4.2.1 PSNR

From Figure (3.4), the comparison of PSNR for both methods is observed for the reconstruction of MR image. Our algorithm significantly removes aliasing and noise noticed in the zero filled, hence providing a better reconstruction. The reconstruction error magnitude of the image for SiFo displays pixel errors of considerably reduced magnitude and fewer structures than that of DLMRI technique by Saiprasad et al.

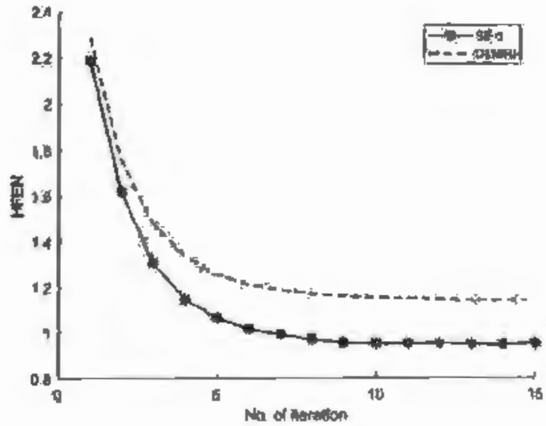
Similarly, the reconstruction of the phantom image has shown the same improved result as the reconstruction of the brain MR image.

3.4.2.2 HFEN

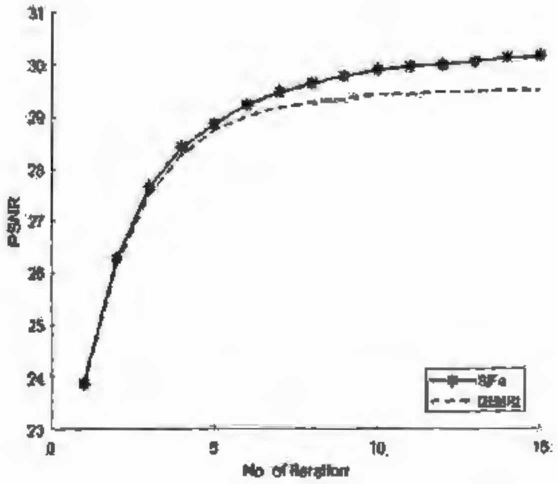
The convergence rate has been observed efficient. If we compare the SiFo with DLMRI, the SiFo converges at the rate 0.95 whereas DLMRI stops at 1.14 (in the reconstruction of brain image) at the end of the executed number of iteration 15. This shows the proposed algorithm outperforms DLMRI regarding reasonable noise. So image features for the reconstruction of the brain and phantom images are smooth, clear and free from the effect of aliasing and artifacts.



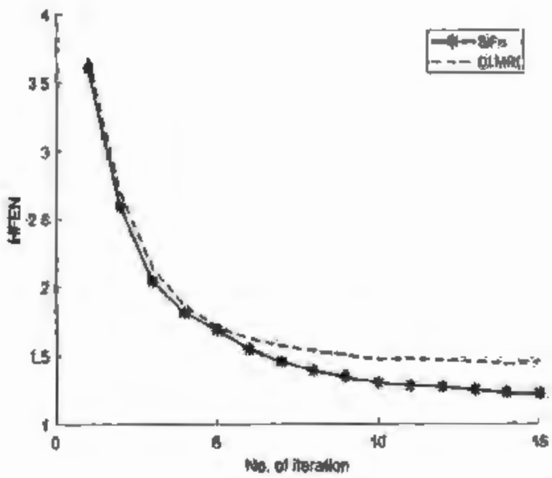
(a)



(b)

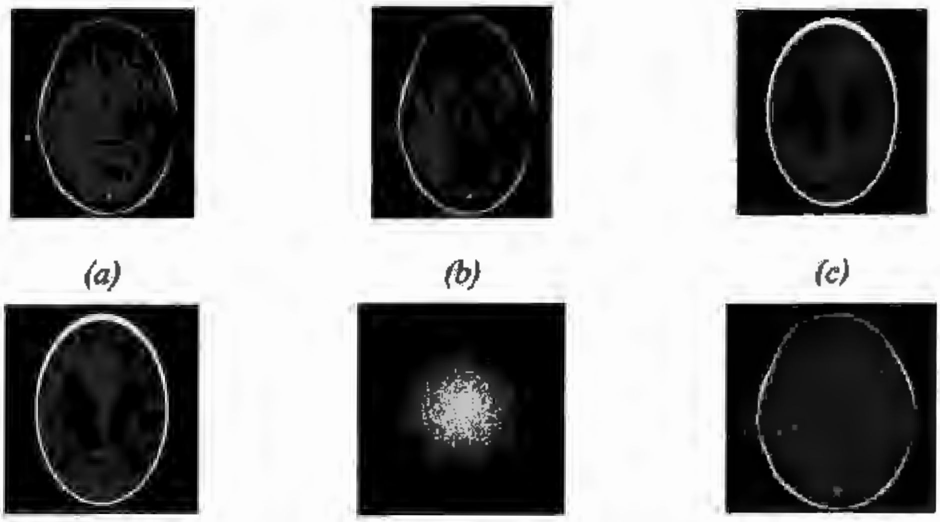


(c)



(d)

Figure 3.4. *Algorithms performance for noisy case. (a) PSNR vs iterations with comparison to DLMRI for brain image (b) HFEN vs iterations with comparison to DLMRI for brain image (c) PSNR vs iterations with comparison to DLMRI for phantom image (d) HFEN vs iterations with comparison to DLMRI for phantom image.*



(a) *(b)* *(c)*

(d) *(e)* *(f)*

Figure 3.5 *Images recovery in noisy case. (a) Recovered MR image of brain by SiFo (b) Recovered MR image of brain by DLMRI (c) Recovered MR image of phantom by SiFo (d) Recovered MR image of phantom by DLMRI (e) sampling mask in k-space with 10 fold (f) Reference MR image of brain.*

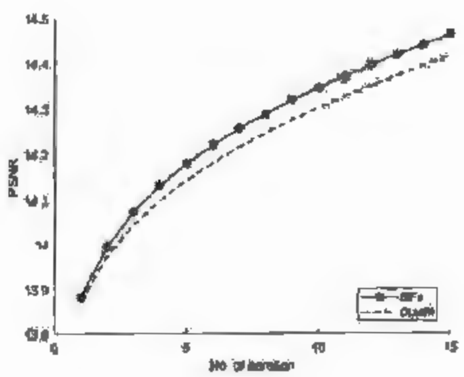
3.4.2.3 Correlation/Similarity index/Sharpness

Although the correlation and similarity index show slight improvement in the reconstruction of the brain and phantom images in noisy case but sharpness indicates noticeable improvement as shown in Table (3.2-3.3). Brightness or sharpness of MRI scan depends on the relaxation time of the specific molecules.

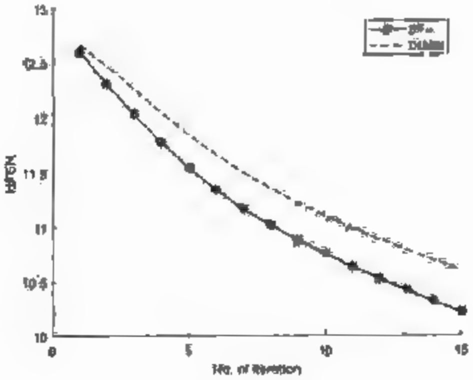
(a) Noisy case for brain Image					
No	Parameters	DLMRI	SiFo	Difference	Remarks
1	**Correlation**	0.9975	0.9981	0.0006	Improved
2	**Similarity Index (SSIM)**	0.8282	0.8902	0.062	Improved
3	**Sharpness**	870.9357	980.2718	109.3361	Improved
(b) Noisy case for phantom Image					
No	Parameters	DLMRI	SiFo	Difference	Remarks
1	**Correlation**	0.9876	0.9892	0.0016	Improved
2	**Similarity Index (SSIM)**	0..7514	0.7716	0.0202	Improved
3	**Sharpness**	3500.2	4177.1	676.9	Improved

Table 3.2. *Performance parameters comparison of SiFo vs. DLMRI algorithms in noisy case for (a) Brain Image (b) Phantom image.*

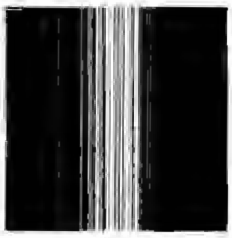
Graphs in Figure (3.6(a, b)) are shown for noisy case with 4 fold Cartesian under sampling for reconstruction of phantom image.



(a)



(b)

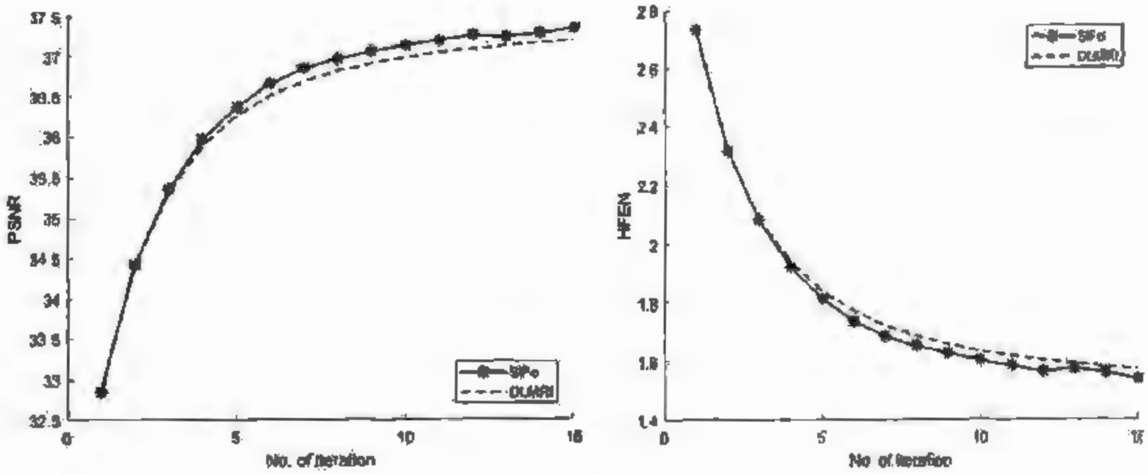


(c)

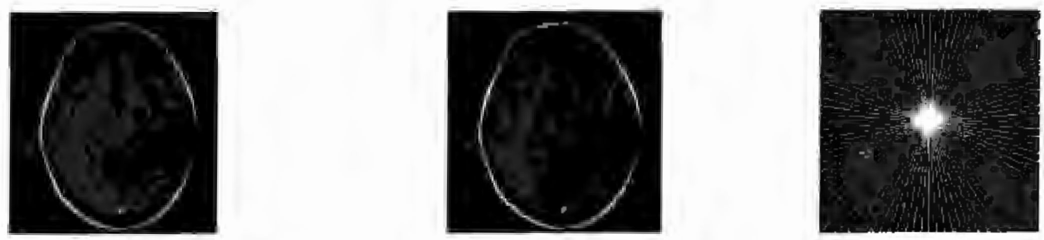
(d)

Figure 3.6*.Algorithms performance in noisy case with Cartesian sampling. (a) PSNR vs iterations with comparison to DLMRI for phantom image (b) HFEN vs iterations with comparison to DLMRI for phantom image (c) Cartesian sampling scheme with 4 fold. (d) Recovered image.*

Performance comparison of SiFo vs DLMRI for noisy brain images for radial sampling are shown in Figure (3.7) and Table (3.3).



(a) *(b)*



(c) *(d)* *(e)*

Figure 3.7. *Algorithms performance of SiFo vs DLMRI in noisy case with radial sampling mask. (a) PSNR vs iterations for phantom image (b) HFEN vs iterations for phantom image (c)Recovered MR image of brain by SiFo (d) Recovered MR image of brain by DLMRI (e) Radial sampling mask in k-space with 6.1 fold undersampling.*

Noisy case for brain image with radial sampling.					
No.	Parameters	DLMRI	SiFo	Difference	Remarks
1	**Correlation**	0.9965	0.9966	0.0001	Improved
2	**Similarity Index (SSIM)**	0.7974	0.8012	0.0038	Improved
3	**Sharpness**	688.437	720.039	31.6017	Improved

Table 3.3. *Performance parameter of Algorithm with noisy case for brain image with radial sampling.*

3.5 Summary

In this chapter, adaptive patch-based dictionary learning framework is proposed by introducing hybrid algorithm of SimCO and FOCUSS for the recovery of MR images. FOCUSS technique has been uses as sparse coding technique, because of its soft constraints, which is considered better for sparse representation of medical images. The pruning process in FOCUSS suppresses the noise during the reconstruction of image. SimCO technique is used for dictionary learning which updates arbitrary number of atoms simultaneously. The problem of increased singular vectors in K-SVD is minimized by regularized version of SimCO.

Chapter 4

MR Image Reconstruction Using Block Dictionary Learning Technique

While representing a class of signals in term of sparsifying transform, it is better to use an adapted learned dictionary instead of using a predefined dictionary as proposed in the recent literature. With this improved method, one can represent the sparsest approximation for the specified signals set. To ease the approximation, atoms of the learned dictionary can further be grouped together to make blocks inside the dictionary that act as a union of small number of subspaces. The block structure of a dictionary can be learned by exploiting the latent structure of the desired signals. Such type of block dictionary leads to block sparse representation of the given signals which can be good for reconstruction of the medical images. In this chapter, a framework is suggested for MRI reconstruction based upon block sparsifying transform (dictionary). Proposed technique develops automatic detection of underlying block structure of MR images given maximum block sizes. It can be achieved by iteratively interchanging between updating block-sparse coding and the block structure of the sparsifying transform (dictionary) of the MR images. Empirically it is shown that block sparse representation performs better for recovery of the given MR image with minimum errors.

4.1 Learning Sparsifying Transform

From a given medical signal $\boldsymbol{x} \in \mathbb{C}^q$ or a patch of the signal extracted as $\boldsymbol{x}_{ij} \in \mathbb{C}^n$, following optimization problem, as discussed in chapter 3 Equation (3.1), may evaluate through the block dictionary learning scheme.

$$\min_{\mathbf{D},\boldsymbol{\theta}} \sum_{ij} \left\|\mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij}\right\|_2^2 \qquad s.t. \left\|\boldsymbol{\theta}_{ij}\right\|_0 \le \tau_0 \ \forall i,j \tag{4.1}$$

Above problem can be expressed by combining all patches $\boldsymbol{x}_{ij}$ column wise in a matrix $\bar{\mathbf{X}} \in \mathbb{C}^{n\times L}$, where L is the total number of patches acquired form of $\boldsymbol{x}$,we get.

$$\min_{\mathbf{D},\boldsymbol{\Theta}} \left\|\bar{\mathbf{X}} - \mathbf{D}\boldsymbol{\Theta}\right\|_F \qquad s.t. \|\boldsymbol{\theta}\|_0 \le \tau_0 \tag{4.2}$$

Where $\boldsymbol{\Theta} \in \mathbb{R}^{K\times L}$ belongs to sparse representation. Optimization formulation used in equations (4.1) and (4.2) are NP hard for the fixed $\mathbf{D}$ and can be solved from many algorithms such as MOD and K-SVD. The CS framework exploits the sparsity of $\boldsymbol{\theta}$ in order to facilitate recovery. With proper chosen $\mathbf{D}$, recovery is possible irrespectively of the location of the nonzero values of $\boldsymbol{\theta}$. This outcome has caused to generate a lot of recovery algorithms. A lot of recent work has been done to find an adaptive structure (block) dictionary for an impressive signal reconstruction.

4.2 Block Dictionary Learning

Block structured dictionaries are learned to utilize any embedded structure in order to produce a more effective sparse representation. Block-structured dictionary method involves two steps: update the dictionary block and then finding the sparse coefficients according to block structure in the dictionary. BOMP (instead of OMP) algorithm is used for updating the sparse representation matrix and BLKSVD [29] to learn the dictionary. Algorithm BOMP chooses the blocks of dictionary sequentially that are best suited to the input signals. The key characteristic of BLKSVD is to update the atoms in blocks and corresponding non-zero coefficients simultaneously.

For a given set of MR signals, we want to find the dictionary whose atoms are sorted into blocks and provide the most accurate representation vectors. An index number is designated to each block and $\boldsymbol{b} \in \mathbb{R}^K$ denotes the vector of block assignments, assigned to the atoms of $\mathbf{D}$. In other words, $\boldsymbol{b}[k]$ denotes the index of a block vector $\mathbf{D}_k$. A vector $\boldsymbol{\theta} \in \mathbb{C}^K$ is s-block sparse over $\boldsymbol{b}$ if its nonzero values are focused only to s blocks. It can be expressed as:

$$\|\boldsymbol{\theta}\|_{0,b} = s \tag{4.3}$$

Where $\|\boldsymbol{\theta}\|_{0,b}$ is the l_0 norm over block $\boldsymbol{b}$ and computes the number of non-zero blocks as defined by $\boldsymbol{b}$. Our objective is to learn the block dictionary $\mathbf{D}$ along with its block structure $\boldsymbol{b}$ having a maximum block size of s that leads to optimal τ_0-block sparse representation $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_p\}_{p=1}^{p=L}$. For block-dictionary learning, Equation (4.2) can be expressed as,

$$\min_{\mathbf{D},\boldsymbol{\Theta}} \|\tilde{\mathbf{X}} - \mathbf{D}\boldsymbol{\Theta}\|_F \quad s.t. \|\boldsymbol{\Theta}_p\|_{0,b} \leq \tau_0 \ , p = 1 \dots . L, \qquad |b_j| \leq s \qquad , j \in [1, N] \tag{4.4}$$

Where N is indicates the number of blocks and b_j is the set of indices represents the list of dictionary atoms in block j and can be expressed as

$$b_j = \{k \in 1,2,3, \dots \dots \dots K | b[k] = j\} \tag{4.5}$$

4.3 Problem Formulation

Reconstructed compressively sampled biomedical MR images typically suffer from numerous artifacts during under sampling of k-space and noise in samples. These are two main causes of artifacts. So a decent dictionary must be capable of minimizing the artifacts which are noticed into zero filled Fourier reconstruction and be consistent to produce reconstructed images by available k-space data. In MRI, the available data is in k-space rather than in an image domain. Proposed formulation has a capability of both designing an adaptive dictionary learning, and also using it to reconstruct the underlying MR image. This is achieved by means of only the under-sampled k-space measurement, $\boldsymbol{z}$. The problem formulation, for reconstructing the MR image $\boldsymbol{x}$, is then becomes:

$$\min_{\mathbf{D},\boldsymbol{\theta},x} \sum_{ij} \left\|\boldsymbol{x}_{ij} - \mathbf{D}\boldsymbol{\theta}_{ij}\right\|_2^2 + \eta \left\|\boldsymbol{\Phi}_u \boldsymbol{x} - \boldsymbol{z}\right\|_2^2 \qquad s.t. \left\|\boldsymbol{\theta}_{ij}\right\|_{0,b} \leq \tau_0 \quad \forall\, i,j \tag{4.6}$$

Cost function of Equation (4.6) is almost same as that discussed in chapter 3 with some more (block) constraints.

4.4 MR Image Reconstruction

Adaptive learning of sparsifying transform in Equation (4.6) is a non-convex problem and is computationally expensive. Typically the problem in Equation (4.6) is solved in two steps. (i) dictionary learning and sparse coding are updated alternately keeping the estimated signal $\boldsymbol{x}$ fixed, while in second step (ii) update the estimated MR signal $\boldsymbol{x}$ to satisfy the data fidelity while keeping dictionary and sparse representation fixed.

4.4.1 UPDATING BLOCK DICTIONARY AND SPARSE CODING

Since MR signal $\boldsymbol{x}$ is fixed, the objective function in Equation (4.6) becomes

$$\min_{\mathbf{D},\boldsymbol{\theta}} \sum_{ij} \left\|\boldsymbol{x}_{ij} - \mathbf{D}\boldsymbol{\theta}_{ij}\right\|_2^2 \qquad s.t. \left\|\boldsymbol{\theta}_{ij}\right\|_{0,b} \leq \tau_0 \ \forall\, i,j \tag{4.7}$$

Combining all patches in Equation (4.7) into matrix $\tilde{\mathbf{X}}$ as described in Equation (4.4). Since the learning sparsifying transform process involves two steps, i.e., block dictionary learning **D** and update the sparse coefficient matrix **Θ**, so we formulate the objective function as follows for given block ***b*** for l^{th} iteration:

$$[\mathbf{D}^l, \mathbf{\Theta}^l] = \min_{\mathbf{D},\mathbf{\Theta}} \|\tilde{\mathbf{X}} - \mathbf{D\Theta}\|_F \quad s.t. \|\mathbf{\Theta}_p\|_{0,b^l} \le \tau_0 \ , p = 1 \ldots\ldots L \tag{4.8}$$

Now applying block KSVD (BLKSVD) algorithm to recover the **D** and **Θ** by optimizing Equation (4.8) on given block structure ***b***. At every l^{th}-iteration, we fix dictionary $\mathbf{D}^{l-1}$ in first step and use the BOMP to solve Equation (4.8) which optimizes as

$$\mathbf{\Theta}^l = \min_{\mathbf{\Theta}} \|\tilde{\mathbf{X}} - \mathbf{D}^{l-1}\mathbf{\Theta}\|_F \qquad s.t. \|\mathbf{\Theta}_p\|_{0,b} \le \tau_0 \ , \quad p = 1 \ldots L \tag{4.9}$$

In second step, we obtain $\mathbf{D}^l$ while fixing $\mathbf{\Theta}^l$, b and $\tilde{\mathbf{X}}$.

$$\mathbf{D}^l = \quad \min_{\mathbf{D}} \|\tilde{\mathbf{X}} - \mathbf{D\Theta}^k\|_F \tag{4.10}$$

Motivated by KSVD algorithm, the blocks are updated in sequence in $\mathbf{D}^{l-1}$ along with its relevant non-zero sparse coefficients in $\mathbf{\Theta}^l$. Details for every block $j \in [1, N]$ are discussed as follows:

Let $\mathbf{E}_{\alpha_j}$be the error matrix of the signals $\tilde{\mathbf{X}}_{\alpha_j}$excluding the contribution of thejth block. We express it as follows:

$$\mathbf{E}_{\alpha_j} = \tilde{\mathbf{X}}_{\alpha_j} - \sum_{k \ne j} \mathbf{D}_{b_k} (\mathbf{\Theta}^{b_k})_{\alpha_j} \tag{4.11}$$

Where $\alpha_j$is the set of indices corresponding to columns in sparse matrix $\mathbf{\Theta}^l$that use the atom $\mathbf{D}_k$.The representative error of signal with indices α_j can be defined as follows.

$$\alpha_j \triangleq \left\|\mathbf{E}_{\alpha_j} - \mathbf{D}_{b_k}\mathbf{\Theta}_{\alpha_j}^{b_k}\right\|_F \tag{4.12}$$

By taking the SVD of error matrix $\mathbf{E}_{\alpha_j}$

$$\mathbf{E}_{\alpha_j} = USV^{\mathrm{T}} \tag{4.13}$$

The dictionary and sparse representation matrix is updated as follows:

$$\mathbf{D}_{b_k} = \left[U_1, \dots \dots \dots \dots \dots U_{|b_k|}\right] \quad (4.14)$$

$$\mathbf{\Theta}_{\alpha_j}^{b_j} = \left[S_1^1 V_1, \dots \dots \dots \dots \dots \dots \dots S_{|b_j|}^{|b_j|} V_{|b_j|}\right]^{\mathrm{T}} \quad (4.15)$$

Iteratively, dictionary and its blocks are continuously updated till the convergence or any predefined number of iterations. When block size is one then both BLKSVD and KSVD become identical.

4.4.2 UPDATING THE ESTIMATED IMAGE(S) FOR RECONSTRUCTION

To update the reconstruction MR image $\boldsymbol{x}$, keep the dictionary and the sparse representation constant then the sub-problem for our cost function in Equation (4.6) can also be written as follows:

$$\min_{x} \sum_{ij} \left\|\mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij}\right\|_2^2 \quad + \quad \eta \left\|\mathbf{\Phi}_u \boldsymbol{x} - \boldsymbol{z})\right\|_2^2 \quad (4.16)$$

The formulation in Equation (4.16) is ,exactly same as described in chapter 3 Equation (3.5), the least squares problem and detailed solution is in appendix-A. The solution is same to reconstruct the image $\boldsymbol{x}$.

Algorithm 4.1:

Goal: Reconstruction of undersampled MR image using block dictionary

Input: $\boldsymbol{z}$= training signal(s) in k-space measurements, s for block size, and τ_0 for sparsity

Output: $\boldsymbol{x}$ = Reconstructed MR image

Initialization: $\boldsymbol{x} = \boldsymbol{x}_0 = \mathbf{\Phi}_u^{\mathrm{H}} \boldsymbol{z}$

Main Iteration:

1. Alternately learn sparsifying transform (dictionary) by BLKSVD and BOMP for sparse coding.
2. Update $\hat{x}$: Every pixel value attained by averaging the impact of patches that covering it
3. $\mathcal{N} \leftarrow \mathcal{FFT}(x)$
4. Restore sampled frequency to update the $\mathcal{N}$ as per Equation (3.6)
5. $\hat{x} \leftarrow$ IFFT ($\mathcal{N}$)

A more extensive pseudocode is presented in Appendix-B2

The complete process for the proposed block dictionary learning is described as below in the Figure (4.1).

4.5 Empirical Result and Discussion

The performance of the proposed algorithm is validated with two cases i.e. noiseless and noisy for the reconstruction of MR Image. We have used same real world image single-slice dataset and parameters as used in chapter 3, except σ , τ_0 *or* sparsity = 6 and block size $s = 3$. Different undersampling factors (2.5 folds to 4 folds) and undersampling schemes (like center dense and Cartesian including 2D random sampling) are used in both noiseless and noisy cases. We have compared our reconstructed images with leading DLMRI [26] method which had already outperformed other CSMRI [21] techniques.

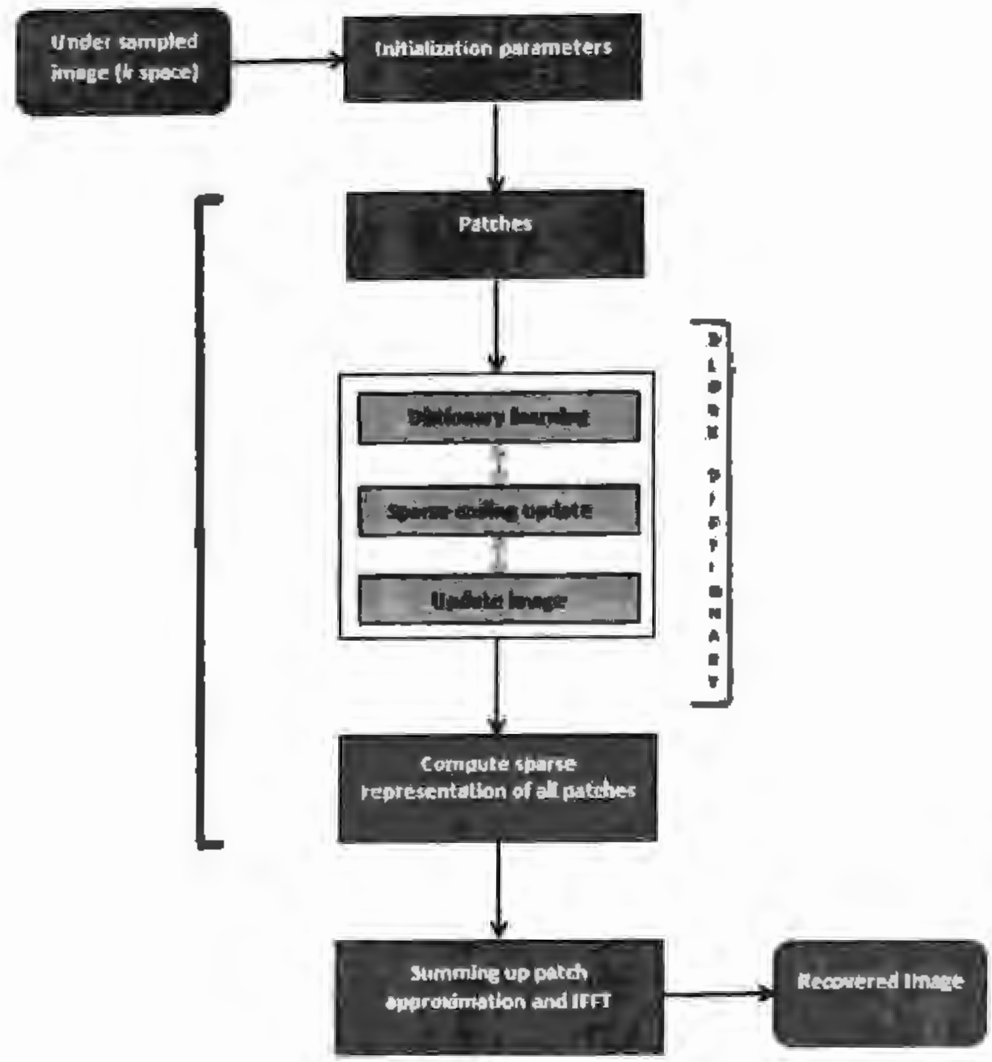


***Figure 4.1** Patch based dictionary learning framework for brain MR image reconstruction based on CS through BLKSVD method*

4.5.1 Performance in Noiseless Scenario:

We performed the proposed method on noiseless scenario first and compared with DLMRI. 2D random center dense and Cartesian sampling schemes with different undersampling factors (2.5 – 4.5) folds are used to reconstruct the image without adding noise in k-space. The dictionary learning scheme by BLKSVD reconstructed the images free from artifacts and aliasing effect in all different sampling schemes. The recovered images are clear with fine edges with fast convergence than DLMRI. In all cases, as shown in Figures (4.2-4.3), our proposed method outperformed the DLMRI in term of PSNR and HFEN.

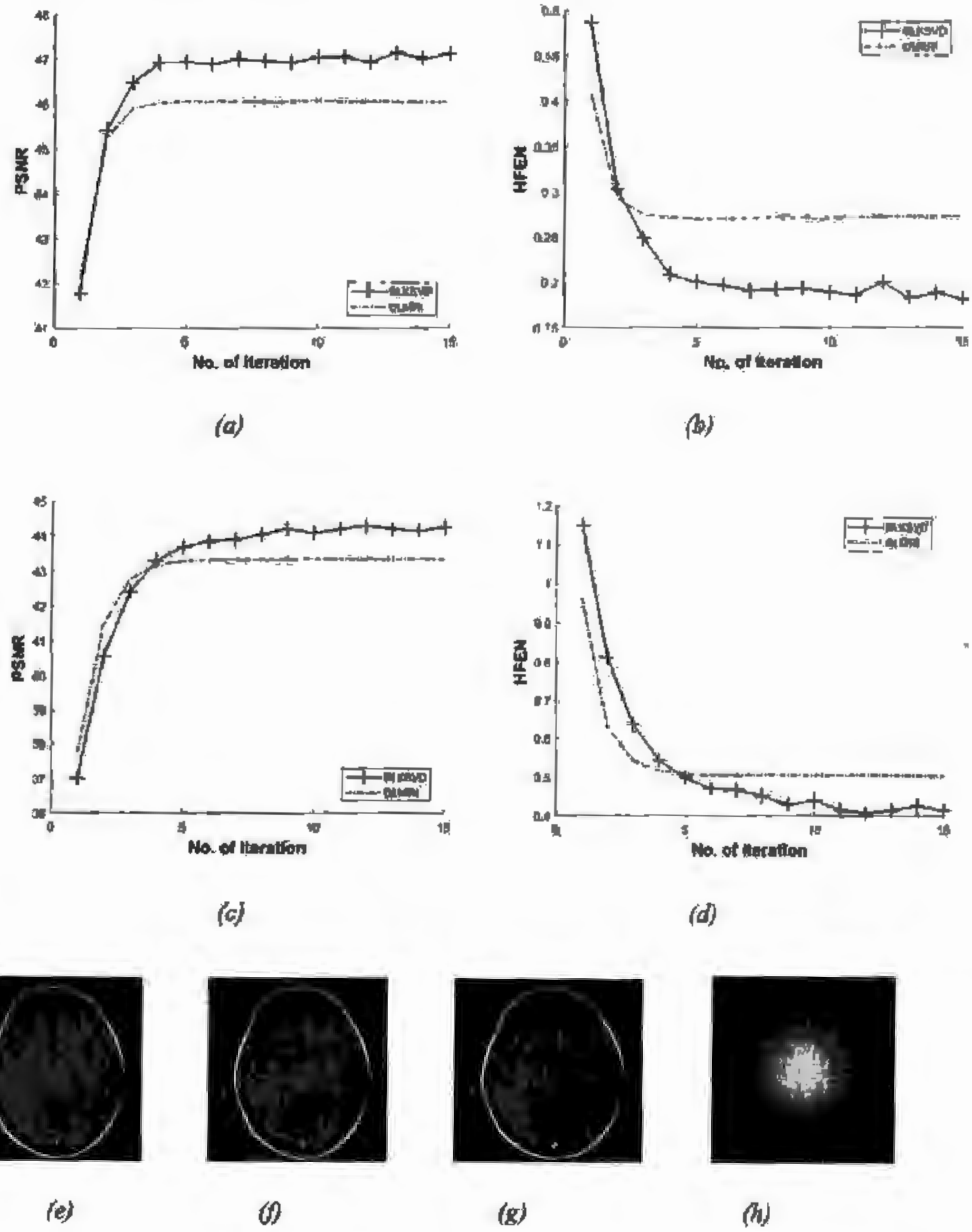


***Figure 4.2**. For Noiseless case: Center dense sampling scheme for reconstruction of brain image (a) PSNR vs iterations with sampling factors 2.5 folds (b) HFEN vs iterations with sampling factors 2.5 folds (c) PSNR vs iterations with sampling factors 4 folds (d) HFEN vs iterations with sampling factors 4 folds. (e) Recovered image with sampling factors 2.5 folds (f) Recovered image with sampling factors 4 fold (g) Reference image (h) Center dense mask.*

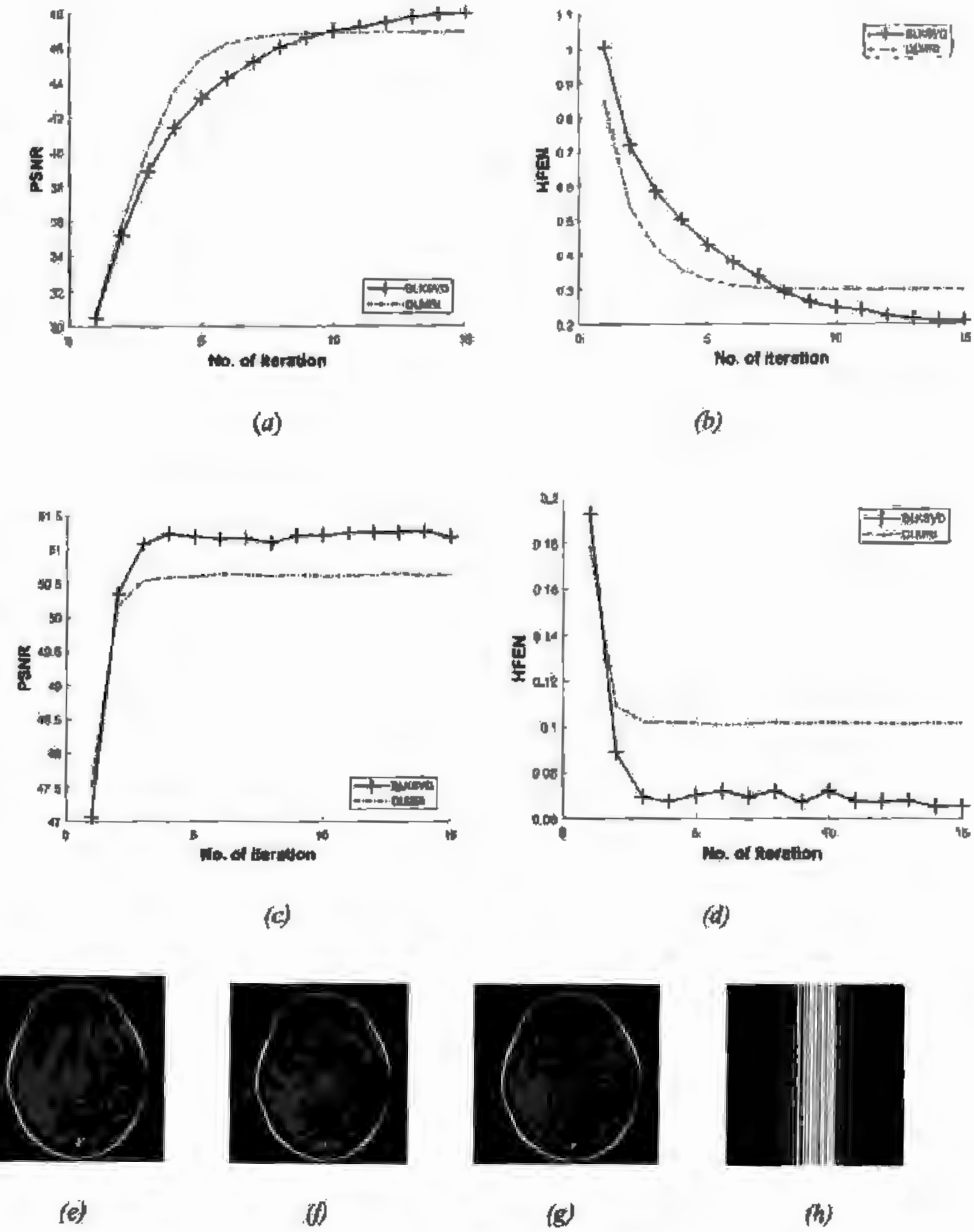


Figure 4.3. *For Noiseless case: Cartesian sampling scheme for reconstruction of brain image (a) PSNR vs iterations with sampling factors 2.5 folds (b) HFEN vs iterations with sampling factors 2.5 folds (c) PSNR vs iterations with sampling factors 4 folds (d) HFEN vs iterations with sampling factors 4 folds. (e) Recovered image with sampling factors 2.5 folds (f) Recovered image with sampling factors 4 folds (g) Reference image (h) Cartesian mask.*

We have also compared DLMRI and the proposed frame work through statistical data with relative difference in the tables for the correlation and similarity index (SSIM). There is slight improvement of correlation and SSIM for center dense mask. It is observed that the correlation in Cartesian schemes on different sampling factors have same results but SSIM has improved as per Table (4.1).

Center Dense Undersampling 2.5 Folds					
No	Parameters	DLMRI	BLKSVD	Difference	Remarks
1	**Correlation**	0.9995	0.9996	0.00010	Improve
2	**Similarity Index (SSIM)**	0.9516	0.9703	0.0187	Improve
Center Dense Undersampling 4 Folds					
No	Parameters	DLMRI	BLKSVD	Difference	Remarks
1	**Correlation**	0.9991	0.9992	0.00010	Improve
2	**Similarity Index (SSIM)**	0.9275	0.9569	0.0294	Improve
Cartesian Undersampling 2.5 Folds					
No	Parameters	DLMRI	BLKSVD	Difference	Remarks
1	**Correlation**	0.9996	0.9996	0.00000	same
2	**Similarity Index (SSIM)**	0.9587	0.9738	0.0151	Improve
Cartesian Undersampling 4 Folds					
No	Parameters	DLMRI	BLKSVD	Difference	Remarks
1	**Correlation**	0.9998	0.9998	0.00000	same
2	**Similarity Index (SSIM)**	0.9836	0.9883	0.0047	Improve

Table 4.1. *Performance parameter (BLKSVD vs DLMRI) of algorithm with noiseless case for brain image*

4.5.2 Performance in Noisy Scenario:

Experiments are performed by adding zero mean white Gaussian noise of standard deviation $\sigma = 3$ in k-space in all cases of center dense and Cartesian sampling at 2.5 fold undersampling. Our proposed method performed better as shown in Figure (4.3) than that of DLMRI in noisy case keeping same parameters and reference brain image and undersampled mask as in noiseless case.

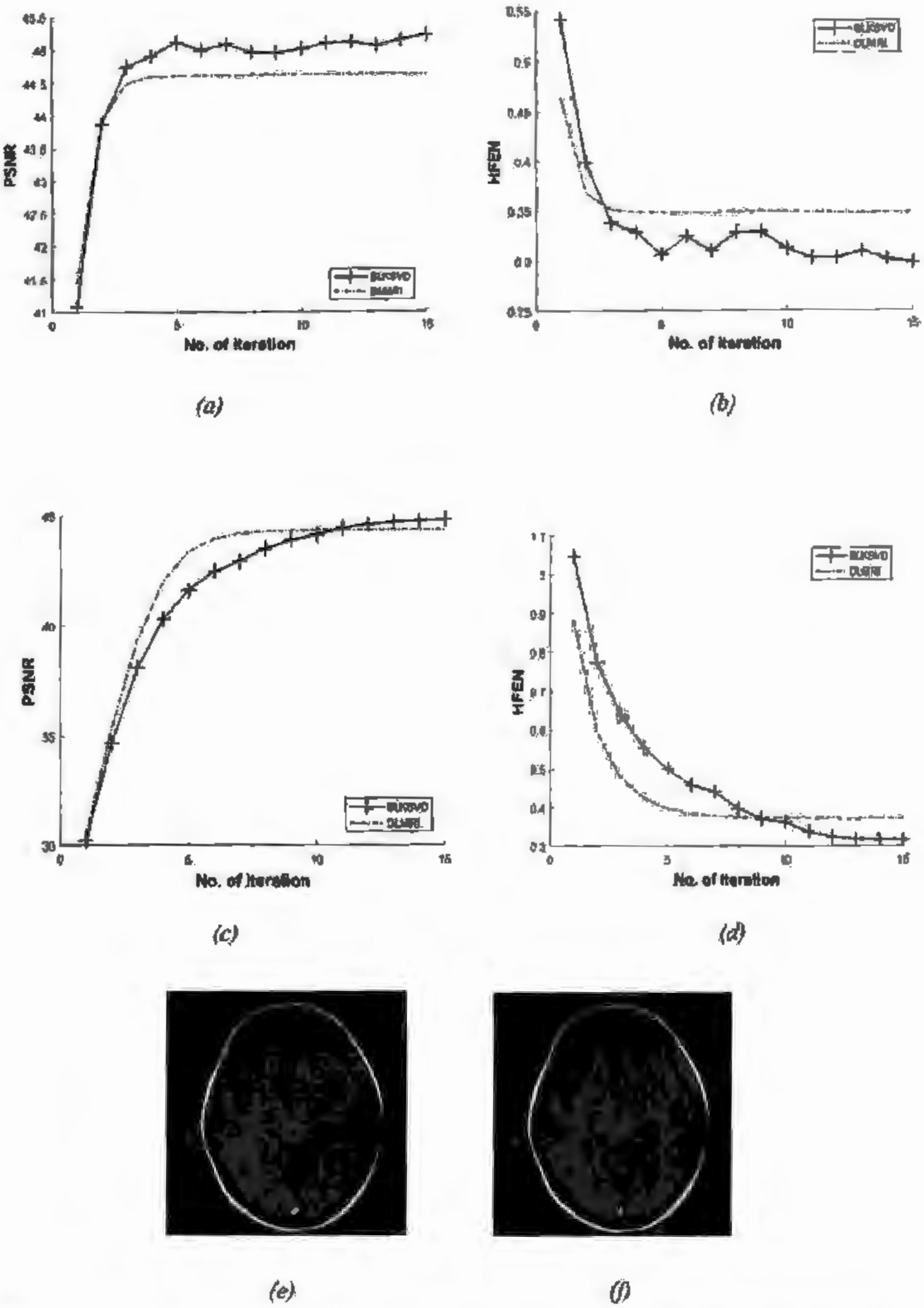


Figure 4.4*.For Noisy case: Reconstruction of brain image (a) PSNR vs iterations with center dense sampling scheme at sampling factors 2.5 folds (b) HFEN vs iterations with center dense sampling scheme at sampling factors 2.5 folds (c) PSNR vs iterations with Cartesian sampling scheme at sampling factors 2.5 folds (d) HFEN vs iterations with Cartesian sampling scheme at sampling factors 2.5 folds. (e) Recovered image center dense sampling scheme at sampling factors 2.5 folds (f) Recovered image center dense sampling scheme at sampling factors 2.5 folds*

The quantitative measurement of SIMM in noisy case has shown slight improvement but correlation is same on both cases of centered dense and Cartesian sampling schemes as observed in Table (4.2).

Center Dense Undersampling 2.5 Folds (Noisy)					
No.	Parameters	DLMRI	BKSVD	Difference	Remarks
1	**Correlation**	0.9994	0.9994	0	same
2	**Similarity Index (SSIM)**	0.9208	0.9364	0.0156	improved
Cartesian Undersampling 2.5 Folds (Noisy)					
No.	Parameters	DLMRI	BKSVD	Difference	Remarks
1	**Correlation**	0.9994	0.9994	0	same
2	**Similarity Index (SSIM)**	0.9068	0.9175	0.0107	improved

Table 4.2. *Performance parameter (BLKSVD vs DLMRI) of algorithm with noisy case for brain image*

4.6 Summary

In this chapter, adaptive patch-based block-structured dictionary learning framework has been introduced for reconstruction of MR images. In this proposed method, BOMP chooses the dictionary blocks sequentially best suited to the applied signals. The key characteristic of BLKSVD is to update the atoms in blocks and corresponding non-zero coefficients simultaneously. So the proposed method combines the advantage of block structure with patch-based adaptive dictionary learning make it possible for the enhanced MR image reconstruction

Chapter 5

Conclusions and Future Work

5.1 Conclusion

Quality of reconstructed (under-sampled) image mainly depends upon sparsifying transform. In this dissertation, different methods are studied for the data-driven adaptation of learning sparsifying transforms for the reconstruction of medical image. Various conclusions and future directions are explained as follows.

In chapter 3, the proposed hybrid formulations of SimCo and FOCUSS (so called SiFo) contribute to represent the given image as sparse in well-conditioned transforms than existing learning transforms typically improved PSNR and HFEN. Soft constraints ability of FOCUSS algorithm has shown better results for sparse representation of medical images. While the pruning process of FOCUSS has played a vital role to suppress the noise during the reconstruction of image because aliasing artifacts and noises are usually isolated. This is the main reason of our proposed (SiFo) scheme which performs well particularly in noisy case. Instead of updating dictionary atoms one by one, SimCO method simultaneously updates an arbitrary number of atoms. This leads to faster learning and computations with the sparse transform. This adapted sparse transform causes to reduce the storage requirement and generalize better than existing learning sparsifying transform. Moreover, problem of increased singular vectors in K-SVD can be minimized by regularized version of SimCO. So Fast convergence with almost accurate reconstruction at high under sampling rates is achieved by the proposed algorithm.

Proposed methods for reconstruction of images are based on patched based learning sparsifying transform instead as whole. Patch-based sparsity dictionary learning has a tendency to capture the local image features effectively and have a potential to eliminate the aliasing artifact without compromising the resolution. This adaptive patch-based dictionary learns from a small number of k-space samples and produces promising reconstruction.

In chapter 4, unique sparsifying transform learning based on block structure is used for general compressed sensing. This framework exploits the transform domain sparsity of overlapping image patches in 2D into block structure, which involves into well organized and efficient update steps. Importantly, BLKSVD is introduced on patch-based dictionary learning in our framework for reconstruction of MR image. The shift from the global image sparsity to patch-based sparsity is added feature in block-structured dictionary learning which has shown better results in MRI reconstruction. Because block patch-based dictionaries has a tendency to capture the local image features efficiently. The use of block-structured dictionary learning on patch based MR image has generated an added averaging effect that eliminates the noise. Moreover, a single MR image can be decomposed into many overlapping patches to train a sparsifying transforms. It suggests a unique framework for simultaneously learning the transform and reconstructing the MR image from under-sampled k-space data.

We have initialized the dictionary in all proposed methods for by extracting left singular vectors from the training data for block dictionary learning in our framework and then normalized each atom of the dictionary. It is observed that by initializing dictionary in this form, convergence becomes faster as compared with randomly initialized dictionary

5.2 Future Work

Future directions for extending the work in this dissertation are as follows

1. The proposed techniques are only applied on MRI so these proposed frame work may also apply for other medical imaging modalities not only for static single slice image but also can be extended to dynamic MRI like fMRI, cardiac MRI etc.
2. It looks promising to explore the application of proposed methodology for other types of noise or artifacts that come during scanning time like head movement etc.
3. BOMP algorithm may be replaced with other time efficient block-sparse approximation methods
4. The values of some parameters such as patch size, sparsity level, dictionary size, error threshold etc., are predefined. If these values do not settle with the ground truth, performance of learning can be degraded by significant level. So it can be solved by developing a non-parametric.

Appendix A

To update the reconstruction image $\boldsymbol{x}$, keep the dictionary and the sparse representation constant then the sub-problem for our cost function in Equation (3.3) become as follows:

$$\min_{\boldsymbol{x}} \sum_{ij} \left\| \mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij} \right\|_2^2 + \eta \left\| \boldsymbol{\Phi}_{\mathrm{u}}\boldsymbol{x} - \boldsymbol{z} \right\|_2^2 \qquad \text{(i)}$$

Differentiate Equation (i) w.r.t $\boldsymbol{x}$ and equate to zero, $\mathbf{W} \in \mathbb{R}$ and $\boldsymbol{\Phi}_{\mathrm{u}} \in \mathbb{C}$

$$\left(\frac{\partial}{\partial x}\right)\left(\sum_{ij} \left\| \mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij} \right\|_2^2 + \eta \left\| \boldsymbol{\Phi}_{\mathrm{u}}\boldsymbol{x} - \boldsymbol{z} \right\|_2^2\right) = 0 \qquad \text{(ii)}$$

$$2\sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \left(\mathbf{W}_{ij}\boldsymbol{x} - \mathbf{D}\boldsymbol{\theta}_{ij}\right) + 2\eta\ \boldsymbol{\Phi}_{\mathrm{u}}^{\mathrm{H}}(\boldsymbol{\Phi}_{\mathrm{u}}\boldsymbol{x} - \boldsymbol{z}) = 0\,, \qquad \text{(iii)}$$

The subscript H and T represents Hermitian transpose operation and real operand respectively. Separating the term belonging to $\boldsymbol{x}$ on left side to find it.

$$\left(\sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \mathbf{W}_{ij} + \eta\ \boldsymbol{\Phi}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{u}}\right)\boldsymbol{x} = \sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \mathbf{D}\boldsymbol{\theta}_{ij} + \eta\ \boldsymbol{\Phi}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{z}\,, \qquad \text{(iv)}$$

In Equation (iv) the 1st term of left side $\sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \mathbf{W}_{ij} \in \mathbb{R}^{q \times q}$ is diagonal matrix and corresponds to image pixel position. These entries are equal to the set of overlapped patches contributing on those pixels. Its diagonal values appear to be all equal and denoted as

$$\sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \mathbf{W}_{ij} = \alpha \mathbf{I}_{\mathrm{q}} \quad \text{where} \quad \mathbf{I}_{\mathrm{q}} \in \mathbb{R}^{q \times q} \qquad \text{(v)}$$

In our case, where diagonal values correspond to place of the image pixel and are equal to the set of overlapped patches contributing to those pixel places.

Putting overlap stride $\mathrm{r} = 1$ for the patches and $\alpha = \mathrm{n}$. When $(\mathrm{f} = \mathrm{q})$, then $\boldsymbol{\Phi} \in \mathbb{C}^{q \times q}$ represent the complete Fourier encoding matrix normalized as $\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi} = \mathbf{I}_{\mathrm{q}}$. Now $\boldsymbol{\Phi x}$ denote full rank data for k-space and putting in Equation (iv) we get,

$$\left(\sum_{ij} \boldsymbol{\Phi}\mathbf{W}_{ij}^{\mathrm{T}} \mathbf{W}_{ij}\boldsymbol{\Phi}^{\mathrm{H}} + \eta\boldsymbol{\Phi}\ \boldsymbol{\Phi}_{\mathrm{u}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{u}}\boldsymbol{\Phi}^{\mathrm{H}}\right)\boldsymbol{\Phi x} = \boldsymbol{\Phi}\sum_{ij} \mathbf{W}_{ij}^{\mathrm{T}} \mathbf{D}\boldsymbol{\theta}_{ij} + \eta\boldsymbol{\Phi}\ \boldsymbol{\Phi}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{z}\,, \qquad \text{(vi)}$$

In Equation (vi), the 2nd term on left side represents the diagonal matrix with scaling η times comprising of zeros and ones. All those ones at diagonal values are related to sampled position k-space. $\mathbf{\Phi}\,\mathbf{\Phi}_u^H\mathbf{z}$ denotes the zero filled Fourier measurements vector and

$$\mathbf{\Phi}\sum_{ij}\ \mathbf{W}_{ij}^{T}\,\mathbf{W}_{ij}\mathbf{\Phi}^{H} = \alpha\mathbf{I}_{q}, \qquad \text{(vii)}$$

Then Equation (vi) becomes

$$\left(\alpha\mathbf{I}_{q} + \mathbf{I}\right)\mathbf{\Phi}\boldsymbol{x} = \alpha\boldsymbol{\mathcal{N}} + \eta\boldsymbol{\mathcal{N}}_{0}, \qquad \text{(viii)}$$

Where:

$$\boldsymbol{\mathcal{N}} \triangleq \mathbf{\Phi}\sum_{ij}\ \mathbf{W}_{ij}^{T}\,\mathbf{D}\,\boldsymbol{\theta}_{ij}\frac{1}{\alpha}, \quad \text{and} \qquad \boldsymbol{\mathcal{N}}_{0} \triangleq \mathbf{\Phi}\ \mathbf{\Phi}_{u}^{H}\,\boldsymbol{z} \qquad \text{(ix)}$$

Dividing Equation (viii) by α both sides we get

$$\left(\mathbf{I}_{q} + \left(\frac{1}{\alpha}\right)\right)\mathbf{\Phi}\boldsymbol{x} = \boldsymbol{\mathcal{N}} + \left(\frac{\eta}{\alpha}\right)\boldsymbol{\mathcal{N}}_{0} \qquad \text{(x)}$$

Here absorbing $\left(\frac{\eta}{\alpha}\right)$ is constant into η then solution of Equation (vi) is as follows:

$$\mathbf{\Phi}\boldsymbol{x}(\mathrm{k_x}, \mathrm{k_y}) = \begin{cases} \boldsymbol{\mathcal{N}}(\mathrm{k_x}, \mathrm{k_y}) & , (\mathrm{k_x}, \mathrm{k_y}) \notin \mho \\ \dfrac{\boldsymbol{\mathcal{N}}(\mathrm{k_x}, \mathrm{k_y}) + \eta\boldsymbol{\mathcal{N}}_{0}(\mathrm{k_x}, \mathrm{k_y})}{1+\eta} & , (\mathrm{k_x}, \mathrm{k_y}) \in \mho \end{cases} \qquad \text{(xi)}$$

Appendix B1

Algorithm Pseudocode: Improved reconstruction of MR scanned images by using dictionary learning scheme.

Inputs: Noisy or noiseless MR Image

Output: Reconstructed MR image from under-sampled data

```
SiFo Parameters Initialization,
x =FFT (Input MR Image),
Add noise in k-space in noisy case,
Applying under-sampling mask,
While number of maximum iterations:
        For i=1: n do
            Create image patches
                For j=1: m do
                    Learn Dictionary from patches
                    Learn sparse codes from patches
                End
            Computing sparse representations of all patches
            Summing up the patch approximation
        End
      x _hat = Unmasked and Inverse FFT of K-Space
     Compute various performances metric (PSNR/HFEN etc.)
End
```

Appendix B2

Algorithm Pseudocode: Enhancing MR image reconstruction using block dictionary

Inputs: Noisy or noiseless MR Image

Output: Reconstructed MR image from under-sampled data

BLKSVD Parameters Initialization,

$\boldsymbol{x}$ =FFT (Input MR Image),

Add noise in *k*-space in noisy case,

Applying under-sampling mask,

While number of maximum iterations:

For i=1: n **do**

Create image patches

For j=1: m **do**

Learn block dictionary

Learn block sparse codes

End

Computing sparse representations of all patches

Summing up the patch approximation

End

$\hat{\boldsymbol{x}}$ = Unmasked and Inverse FFT of *k*-Space

Compute various performances metric

End

REFERENCES

[1] yan yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for Compressive Sensing MRI," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 10–18.

[2] R. W. Liu, W. Yin, L. Shi, J. Duan, S. C. H. Yu, and D. Wang, "Undersampled CS image reconstruction using nonconvex nonsmooth mixed constraints," *Multimed. Tools Appl.*, vol. 78, no. 10, pp. 12749–12782, May 2019.

[3] B. Deka and S. Datta, *Compressed Sensing Magnetic Resonance Image Reconstruction Algorithms*, vol. 9. Singapore: Springer Singapore, 2019.

[4] J. Cao, S. Liu, H. Liu, and H. Lu, "CS-MRI reconstruction based on analysis dictionary learning and manifold structure regularization," *Neural Networks*, vol. 123, pp. 217–233, Mar. 2020.

[5] Y. Heiser, Y. Oiknine, and A. Stern, "Compressive hyperspectral image reconstruction with deep neural networks," in *Big Data: Learning, Analytics, and Applications*, 2019, p. 21.

[6] J. Cao, S. Liu, H. Liu, X. Tan, and X. Zhou, "Sparse representation of classified patches for CS-MRI reconstruction," *Neurocomputing*, vol. 339, pp. 255–269, Apr. 2019.

[7] S. Lachner *et al.*, "Compressed sensing reconstruction of 7 Tesla 23Na multi-channel breast data using 1H MRI constraint," *Magn. Reson. Imaging*, vol. 60, pp. 145–156, Jul. 2019.

[8] H. Zhao, Y. Liu, C. Huang, and T. Wang, "Hybrid-Weighted Total Variation and Nonlocal Low-Rank-Based Image Compressed Sensing Reconstruction," *IEEE Access*, vol. 8, pp. 23002–23010, 2020.

[9] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.

[10] Y. B.-2008 I. T. and A. Workshop and undefined 2008, "Spectrum-blind sampling and compressive sensing for continuous-index signals," *ieeexplore.ieee.org*.

[11] J. Ye, Y. Bresler, P. M.-I. J. of C. Vision, and U. 2002, "A self-referencing level-set method for image reconstruction from sparse Fourier samples," *Springer*.

[12] M. Gastpar, Y. B.-I. C. on Acoustics, U. Speech, and U. 2000, "GastparB00.pdf," *infoscience.epfl.ch*.

[13] Y. Bresler, … M. G.-P. 1999 I., and undefined 1999, "Image compression on-the-fly by universal sampling in Fourier imaging systems," *raman.freeshell.org*.

[14] R. Venkataramani and Y. Bresler, "Further results on spectrum blind sampling of 2D signals," in *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)*, vol. 2, pp. 752–756.

[15] P. Feng, "Universal spectrum blind minimum rate sampling and reconstruction of multiband signals," 1997.

[16] Ping Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, vol. 3, pp. 1688–1691.

[17] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008.

[18] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[19] M. Lustig, J. Santos, J. Lee, and D. Donoho, "Application of compressed sensing for rapid MR imaging," *SPARS,(Rennes*, no. c, pp. 3–5, 2005.

[20] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed Sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.

[21] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of oompressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.

[22] C. Qiu, W. Lu, N. V.- ICASSP, and undefined 2009, "Real-time dynamic MR image reconstruction using Kalman Filtered Compressed Sensing.," *eng.iastate.edu*.

[23] Y. Kim, M. S. Nadar, and A. Bilgin, "Wavelet-Based Compressed Sensing Using a Gaussian Scale Mixture Model," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3102–3108, Jun. 2012.

[24] C. Jiang, Q. Zhang, R. Fan, and Z. Hu, "Super-resolution CT Image Reconstruction Based on Dictionary Learning and Sparse Representation," *Sci. Rep.*, vol. 8, no. 1, p. 8799, Dec. 2018.

[25] E. Candes, T. T. preprint math/0502327, and undefined 2005, "Decoding by linear programming," *arxiv.org*.

[26] S. Ravishankar and Y. Bresler, "MR Image Reconstruction From Highly

Undersampled k-Space Data by Dictionary Learning," *IEEE Trans. Med. Imaging*, vol. 30, no. 5, pp. 1028–1041, May 2011.

[27] K. Engan, S. Aase, and J. Hakon Husoy, "Method of optimal directions for frame design," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*, 1999, pp. 2443–2446 vol.5.

[28] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[29] Z. Zhang, W. Jiang, Z. Zhang, S. Li, G. Liu, and J. Qin, "Scalable Block-Diagonal Locality-Constrained Projective Dictionary Learning," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 2019, pp. 4376–4382.

[30] K. Engan, K. Skretting, and J. H. Husøy, "Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation," *Digit. Signal Process.*, vol. 17, no. 1, pp. 32–49, Jan. 2007.

[31] K. Skretting and K. Engan, "Recursive Least Squares Dictionary Learning Algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2121–2130, Apr. 2010.

[32] Wei Dai, Tao Xu, and Wenwu Wang, "Simultaneous Codeword Optimization (SimCO) for Dictionary Update and Learning," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6340–6353, Dec. 2012.

[33] W. Dai, T. Xu, and W. Wang, "Dictionary learning and update based on simultaneous codeword optimization (SimCO)," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 2037–2040.

[34] S. G. Mallat and Zhifeng Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, 1993.

[35] J. A. Tropp, "Greed is Good: Algorithmic Results for Sparse Approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[36] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the Reconstruction of Block-Sparse Signals With an Optimal Number of Measurements," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3075–3085, Aug. 2009.

[37] Y. C. Eldar and M. Mishali, "Robust Recovery of Signals From a Structured Union of Subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[38] Y. Eldar, "Robust recovery of signals from a union of subspaces," *Arxiv Prepr. arXiv0807.4581*, vol. 55, no. 1081, pp. 1–29, 2008.

[39] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-Sparse Signals: Uncertainty Relations and Efficient Recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[40] K. Gedalyahu and Y. C. Eldar, "Time-Delay Estimation From Low-Rate Samples: A Union of Subspaces Approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3017–3031, Jun. 2010.

[41] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-Sparse Signals: Uncertainty Relations and Efficient Recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[42] Y. C. Eldar and H. Bolcskei, "Block-sparsity: Coherence and efficient recovery," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 2885–2888.

[43] L. Zelnik-Manor, K. Rosenblum, and Y. C. Eldar, "Dictionary optimization for block-sparse representations," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2386–2395, May 2012.

[44] L. Meier, S. Van De Geer, and P. Bühlmann, "The group lasso for logistic regression," *J. R. Stat. Soc. Ser. B (Statistical Methodol.*, vol. 70, no. 1, pp. 53–71, Jan. 2008.

[45] T. Hesterberg, N. H. Choi, L. Meier, and C. Fraley, "Least angle and ℓ 1 penalized regression: A review * †," *Stat. Surv.*, vol. 2, pp. 61–93, 2008.

[46] J. C. Ye, S. Tak, Y. Han, and H. W. Park, "Projection reconstruction MR imaging using FOCUSS," *Magn. Reson. Med.*, vol. 57, no. 4, pp. 764–775, Apr. 2007.

[47] G. Sreeram, Haris B C, and R. Sinha, "Improved speaker verification using block sparse coding over joint speaker-channel learned dictionary," in *TENCON 2015 - 2015 IEEE Region 10 Conference*, 2015, pp. 1–5.

[48] L. Meier, S. Van De Geer, and P. Bühlmann, "The group lasso for logistic regression," *J. R. Stat. Soc. Ser. B (Statistical Methodol.*, vol. 70, no. 1, pp. 53–71, Jan. 2008.

[49] J. Schlemper, J. Caballero, J. V. Hajnal, A. Price, and D. Rueckert, "A Deep Cascade of Convolutional Neural Networks for MR Image Reconstruction," in *Springer*, 2017, pp. 647–658.

[50] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.*, vol. 63, no. 13, 2018.

[51] J. Schlemper *et al.*, "Stochastic Deep Compressive Sensing for the Reconstruction of Diffusion Tensor Cardiac MRI," 2018, pp. 295–303.

[52] M. Seitzer *et al.*, "Adversarial and perceptual refinement for compressed sensing MRI reconstruction," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11070 LNCS, pp. 232–240.

[53] G. Yang, S. Yu, H. Dong, … G. S.-I. transactions on, and undefined 2017, "DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *ieeexplore.ieee.org*.

[54] X. Fan, Q. Lian, and B. Shi, "Compressed-Sensing MRI Based on Adaptive Tight Frame in Gradient Domain," *Appl. Magn. Reson.*, vol. 49, no. 5, pp. 465–477, May 2018.

[55] M. Yuan, B. Yang, Y. Ma, J. Zhang, R. Zhang, and K. Zhan, "Compressed sensing undersampled MRI reconstruction using iterative shrinkage thresholding based on NSST," in *2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2014, pp. 653–658.

[56] Xi Peng and Dong Liang, "MR Image Reconstruction with Convolutional Characteristic Constraint (CoCCo)," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1184–1188, Aug. 2015.

[57] Y. C. Eldar and M. Mishali, "Robust Recovery of Signals From a Structured Union of Subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[58] S. Mallat, *A Wavelet Tour of Signal Processing*. Elsevier, 2009.

[59] W. Dong, G. Shi, Y. Ma, and X. Li, "Image Restoration via Simultaneous Sparse Coding: Where Structured Sparsity Meets Gaussian Scale Mixture," *Int. J. Comput. Vis.*, vol. 114, no. 2–3, pp. 217–232, Sep. 2015.

[60] M. Zibulevsky and M. Elad, "L1-L2 Optimization in Signal and Image Processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76–88, May 2010.

[61] W. Feng, C. Feng-wei, and W. Jia, "Reconstruction Technique Based on the Theory of Compressed Sensing Satellite Images," *Open Electr. Electron. Eng. J.*, vol. 9, no. 1, pp. 74–81, Mar. 2015.

[62] B. Mailhe, R. Gribonval, F. Bimbot, M. Lemay, P. Vandergheynst, and J.-M. Vesin, "Dictionary learning for the sparse modelling of atrial fibrillation in ECG signals," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 465–468.

[63] I. Tosic, I. Jovanovic, P. Frossard, M. Vetterli, and N. Duric, "Ultrasound

tomography with learned dictionaries," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 5502–5505.

[64] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order methoda)," *Med. Phys.*, vol. 37, no. 9, pp. 5113–5125, Aug. 2010.

[65] I. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.

[66] R. Baraniuk, "Compressive Sensing [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[67] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[68] J. A. Tropp, "Greed is Good: Algorithmic Results for Sparse Approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[69] D. Needell and R. Vershynin, "Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit," *Found. Comput. Math.*, vol. 9, no. 3, pp. 317–334, Jun. 2009.

[70] D. Needell and R. Vershynin, "Signal Recovery From Incomplete and Inaccurate Measurements Via Regularized Orthogonal Matching Pursuit," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 2, pp. 310–316, Apr. 2010.

[71] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Commun. ACM*, vol. 53, no. 12, p. 93, Mar. 2008.

[72] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *J. R. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, Jan. 1996.

[73] D. Needell, J. Tropp, and R. Vershynin, "Greedy signal recovery review," in *2008 42nd Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 1048–1050.

[74] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[75] M. Elad, B. Matalon, J. Shtok, and M. Zibulevsky, "A wide-angle view at iterated shrinkage algorithms," 2007, p. 670102.

[76] D. Wipf and B. D. Rao, "Sparse Bayesian Learning for Basis Selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[77] G. H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "Fast Sparse Representation Based on Smoothed ℓ0 Norm," in *Independent Component Analysis and Signal Separation*, vol. 4666 LNCS, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 389–396.

[78] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Natl. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

[79] D. Baron, S. Sarvotham, and R. G. Baraniuk, "Bayesian Compressive Sensing Via Belief Propagation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 269–280, Jan. 2010.

[80] L. Wang *et al.*, "Particle Swarm Optimization based dictionary learning for remote sensing big data," *Knowledge-Based Syst.*, vol. 79, pp. 43–50, May 2015.

[81] H. Haider, J. A. Shah, S. Ikram, and I. A. Latif, "Sparse signal recovery from compressed measurements using hybrid particle swarm optimization," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 2017, pp. 429–433.

[82] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[83] K. P. Pruessmann, "Encoding and reconstruction in parallel MRI," *NMR Biomed.*, vol. 19, no. 3, pp. 288–299, May 2006.

[84] A. Deshmane, V. Gulani, M. A. Griswold, and N. Seiberlich, "Parallel MR imaging," *Journal of Magnetic Resonance Imaging*, vol. 36, no. 1. pp. 55–72, Jul-2012.

[85] L. Grady, 692,549 JR Polimeni - US Patent 8, and undefined 2014, "Method for reconstructing images of an imaged subject from a parallel MRI acquisition," *Google Patents*.

[86] M. Uecker *et al.*, "ESPIRiT-an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 990–1001, Mar. 2014.

[87] A. Deshmane, V. Gulani, M. A. Griswold, and N. Seiberlich, "Parallel MR imaging," *J. Magn. Reson. Imaging*, vol. 36, no. 1, pp. 55–72, Jul. 2012.

[88] M. Elad and M. Aharon, "Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[89] R. Chartrand, "Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data," in *2009 IEEE International Symposium on*

Biomedical Imaging: From Nano to Macro, 2009, pp. 262–265.

[90] X. Qu *et al.*, "Undersampled MRI reconstruction with patch-based directional wavelets," *Magn. Reson. Imaging*, vol. 30, no. 7, pp. 964–977, Sep. 2012.

[91] M. Hong, Y. Yu, H. Wang, F. Liu, and S. Crozier, "Compressed sensing MRI with singular value decomposition-based sparsity basis," *Phys. Med. Biol.*, vol. 56, no. 19, pp. 6311–6325, Oct. 2011.

[92] M. Protter and M. Elad, "Image Sequence Denoising via Sparse and Redundant Representations," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 27–35, Jan. 2009.

[93] J. Mairal, M. Elad, and G. Sapiro, "Sparse Representation for Color Image Restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[94] M. Aharon and M. Elad, "Sparse and Redundant Modeling of Image Content Using an Image-Signature-Dictionary," *SIAM J. Imaging Sci.*, vol. 1, no. 3, pp. 228–247, Jan. 2008.

[95] J. Mairal, G. Sapiro, and M. Elad, "Learning Multiscale Sparse Representations for Image and Video Restoration," *Multiscale Model. Simul.*, vol. 7, no. 1, pp. 214–241, Jan. 2008.

[96] G. Yu, G. Sapiro, and S. Mallat, "Image modeling and enhancement via structured sparse model selection," in *2010 IEEE International Conference on Image Processing*, 2010, pp. 1641–1644.

[97] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double Sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, Mar. 2010.

[98] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 3501–3508.

[99] H. Y. Liao and G. Sapiro, "Sparse representations for limited data tomography," in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2008, pp. 1375–1378.

[100] M. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[101] E. J. Candès and D. L. Donoho, "Ridgelets: a key to higher-dimensional intermittency?," *Philos. Trans. R. Soc. London. Ser. A Math. Phys. Eng. Sci.*, vol. 357, no. 1760, pp. 2495–2509, Sep. 1999.

[102] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[103] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Probl.*, vol. 23, no. 3, pp. 947–968, Jun. 2007.

[104] Y. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pp. 40–44.

[105] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[106] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *Electroencephalogr. Clin. Neurophysiol.*, vol. 95, no. 4, pp. 231–251, Oct. 1995.

[107] R. Tibshirani, I. Johnstone, T. Hastie, and B. Efron, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, Apr. 2004.

[108] R. Rubinstein, T. Faktor, and M. Elad, "K-SVD dictionary-learning for the analysis sparse model," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 5405–5408.

[109] M. Yaghoobi, S. Nam, R. Gribonval, and M. E. Davies, "Noise aware analysis operator learning for approximately cosparse signals," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 5409–5412.

[110] A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," 2004.

[111] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 1, pp. 59–73, Jul. 2011.

[112] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies, "Greedy-like algorithms for the cosparse analysis model," *Linear Algebra Appl.*, vol. 441, pp. 22–60, Jan. 2014.

[113] Y. Liu, T. Mi, and S. Li, "Compressed Sensing With General Frames via Optimal-Dual-Based $\ell _{1}$-Analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4201–4214, Jul. 2012.

[114] B. Ophir, M. Elad, ... N. B.-2011 19th E., and U. 2011, "Sequential minimal

eigenvalues-an approach to analysis dictionary learning," *ieeexplore.ieee.org*, 2012.

[115] "Analysis operator learning for overcomplete cosparse representations - IEEE Conference Publication." [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7074220. [Accessed: 28-Dec-2019].

[116] R. Venkataramani and Y. Bresler, "Further results on spectrum blind sampling of 2D signals," in *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)*, vol. 2, pp. 752–756.

[117] E. Candes and T. Tao, "Decoding by Linear Programming," Feb. 2005.

[118] D. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[119] R. Chartrand, "Exact Reconstruction of Sparse Signals via Nonconvex Minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.

[120] G. Harikumar, C. Couvreur, and Y. Bresler, "Fast optimal and suboptimal algorithms for sparse solutions to linear inverse problems," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, vol. 3, pp. 1877–1880.

[121] G. Harikumar and Y. Bresler, "A new algorithm for computing sparse solutions to linear inverse problems," in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, 1996, vol. 3, pp. 1331–1334.

[122] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *Electroencephalogr. Clin. Neurophysiol.*, vol. 95, no. 4, pp. 231–251, Oct. 1995.

[123] R. Remi and K. Schnass, "Dictionary identification-sparse matrix-factorization via ℓ -minimization," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3523–3539, Jul. 2010.

[124] Y. Chen, X. Ye, and F. Huang, "A novel method and fast algorithm for MR image reconstruction with significantly under-sampled data," *Inverse Probl. Imaging*, vol. 4, no. 2, pp. 223–240, May 2010.

[125] J. Trzasko and A. Manduca, "Highly Undersampled Magnetic Resonance Image Reconstruction via Homotopic l0- Minimization," *IEEE Trans. Med. Imaging*, vol. 28, no. 1, pp. 106–121, Jan. 2009.

[126] A. Jain, *Fundamentals of digital image processing*. 1989.

[127] G. Blanchet and L. Moisan, "An explicit sharpness index related to global phase coherence," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 1065–1068.

[128] A. Bilgin, Y. Kim, F. Liu, M. N.-P. of the 18th, and undefined 2010, "Dictionary design for compressed sensing MRI," *pdfs.semanticscholar.org*.