

Coupled K-SVD Dictionary Learning for Single Image Super-Resolution in Wavelet Domain

Junaid Ahmed

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
August, 2015
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Serhan Çiftçiođlu
Acting Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Chair, Department of Electrical and Electronic
Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hüseyin Özkaramanlı
Supervisor

Examining Committee

1. Prof. Dr. Hasan Demirel _____

2. Prof. Dr. Hüseyin Özkaramanlı _____

3. Assoc. Prof. Dr. Erhan A. İnce _____

ABSTRACT

This thesis introduces coupled K-Singular Value Decomposition (K-SVD) algorithm in wavelet domain for Single Image Super-Resolution (SISR). In the coupled K-SVD the best low-rank approximation given by the SVD is implemented to update the LR and HR dictionaries which in turn help to enforce the equality of the sparse representation coefficients at two resolution levels. Wavelet domain produces better results due to desirable properties such as persistence across scale, compactness, directionality and analysis in many levels with the addition of redundancy in the sparse representations. Using this approach, one can design multiple structured redundant dictionaries, which can potentially help reduce the number of dictionary atoms. Three pairs of coupled low and high resolution wavelet subband dictionaries are designed. Given the low resolution image one first estimates the sparse representation coefficients using the low resolution dictionary and then reconstructs the high resolution image using the calculated low resolution sparse coefficients and high resolution dictionary. This approach generates HR images that are competitive or even better when compared with the state of the art algorithms. Results are improved in terms of PSNR and SSIM in comparison.

Keywords: Coupled K-SVD, Wavelets, Single Image Super-Resolution.

ÖZ

Bu tez tek resimlerin çözünürlüğünü artırmak için dalgacık bölgesinde birleştirilmiş K-SVD algoritmasını tanıtmaktadır. Birleştirilmiş K-SVD algoritmasında en iyi düşük kerte SVD yakınlaştırması kullanılarak hem alçak hem de yüksek çözünürlükteki sözlükler değiştirilmektedir. Böylece, alçak ve yüksek çözünürlükteki seyrek temsiliyet katsayılarının eşitliği zorlanmaktadır. Dalgacık uzayının kullanımı daha iyi sonuçlar vermektedir. Bunun sebebi dalgacıkların ölçeklemeye dayanaklı olması, üç temel yöne cevap verebilmesi ve ayrıtların (edge) değişik dalgacık seviyelerinde iz bırakabilmesidir. Ayrıca, dalgacık dönüşümü tasarlanan sözlüklerde artıklık yaratmaktadır. Sözlükler dalgacıkların yapılarını taşıyacakları için daha küçük olmaları mümkün kılınmaktadır. Böylece hesaplama karmaşıklığının da azaltılması sağlanacaktır. Her bir dalgacık bandı için bir çift ilintili sözlük tasarlanmıştır. Çiftler alçak ise yüksek çözünürlükteki sözlüklerden oluşmaktadır. Verilen bir alçak çözünürlükteki sözlüklerden oluşmaktadır. Verilen bir alçak çözünürlükteki resim önce yamalarına ayrılmakta ve her yama için alçak çözünürlükteki sözlük kullanılarak seyrek temsiliyet katsayıları hesaplanır. Yüksek çözünürlükteki seyrek temsiliyet katsayılarının aynı olduğun varsayımı ile yüksek çözünürlükteki ilintili sözlük kullanılmak suretiyle yüksek çözünürlük yamaları hesaplanmaktadır. Yamalardan örtüştürerek toplama yöntemi ile yüksek çözünürlük resmini oluşturmaktadır. Bu yaklaşım ile elde edilen sonuçlar literatürdeki en iyi algoritmalara kıyasla hem PSNR hem de SSIM ölçütleri bağlamında önemli iyileştirmeler sağlamaktadır.

Anahtar Kelimeler:Coupled K-SVD, Dalgacıklar, Tek Görüntü Süper Çözünürlük

*Dedicated to my parents and family and all those who supported me to complete my
MS research.*

ACKNOWLEDGEMENT

I am very thankful to my parents and family for supporting me. I am really thankful to my supervisor for providing me the skills, guidance and knowledge without which I would not be able to complete my Master.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZ	iv
DEDICATION	v
ACKNOWLEDGEMENT	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS.....	xi
1 INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Motivation.....	2
1.3 Thesis Contributions.....	3
1.4 Thesis Outline.....	5
2 LITERATURE REVIEW	6
2.1 Sparse Model.....	6
2.1.1 Basic Idea.....	6
2.2 Orthogonal Matching Pursuit.....	7
2.3 Iterative Reweighted Least Squares (IRLS).....	8
2.4 Least Angle Regression Stage wise.....	8
2.5 Dictionary Learning.....	10
2.5.1 Method of Optimal Direction.....	11
2.5.2 K-SVD Algorithm.....	11
2.6 Applications.....	11
2.6.1 Inpainting.....	11

2.6.2 Denoising.....	13
2.6.3 Image Compression.....	14
2.6.4 Image Superresolution.....	14
3 SUPERRESOLUTION	16
3.1 Introduction.....	16
3.2 Joint Sparse Coding.....	17
3.3 Coupled Dictionary Learning.....	18
3.4 Coupled K-SVD Algorithm.....	20
4 THE PROPOSED SUPERRESOLUTION METHOD.....	22
4.1 Single Image Superresolution.....	22
4.2 Proposed SR Approach.....	23
4.2.1 The Proposed Dictionary Learning Algorithm.....	24
4.2.2 The Training Image Dataset.....	27
4.2.3 The Proposed Image Reconstruction Algorithm.....	28
5 SIMULATION AND RESULTS.....	30
5.1 Introduction.....	30
5.2 Quantitative Results.....	31
5.3 Qualitative Result.....	33
6 CONCLUSION.....	35
6.1 Conclusion.....	35
6.2 Future Task.....	35
REFERENCES.....	36

LIST OF TABLES

Table 5.1: The PSNR (up) and SSIM (down) value comparison of the proposed algorithm with the Bicubic, Yang et al, Xu et al, Nazzal et al algorithms.....	32
--	----

LIST OF FIGURES

Figure 4.1: HR Dictionaries learned using coupled K-SVD (a) Diagonal Structured (b) Horizontal Structured (c) Vertical Structured.....	25
Figure 4.2: Proposed Dictionary Learning Algorithm	26
Figure 4.3: Samples of images used in dictionary learning.....	26
Figure 4.4: The Proposed Image Reconstruction Algorithm.....	29
Figure 5.1: Visual Comparison of the Image 1 from the Kodak data set with the Bicubic, Yang et al, Xu et al and Nazzal et al algorithms.....	33
Figure 5.2: Visual Comparison of the Boat Image with the Bicubic, Yang et al, Xu et al and Nazzal et al algorithms.....	34

LIST OF ABBREVIATIONS

v_k	k^{th} right singular value
u_k	k^{th} left singular value
φ	Down Sampling and Blurring Operator
σ_k	Singular values
σ_{XX}	Covariance values
\dagger	Pseudo Inverse
γ_i	Sparse Representation
d_r	r^{th} atom of dictionary
μ_X	Average of X
$\ \cdot\ _2$	Euclidean Norm
$\ \cdot\ _0$	Zero Norm
$\ \cdot\ _F$	Frobenious Norm
DWT	Discrete Wavelet Transform
HR	High Resolution
IRLS	Iterative Reweighted Least Squares
KSVD	K-Mean Singular Value Decomposition
LASSO	Least Absolute Shrinkage and Selection Operator
LARS	Least Angle Regression Stage-wise
LR	Low Resolution
MOD	Method of Optimal Directions
MSE	Mean Squared Error
OMP	Orthogonal Matching Pursuit
QCQP	Quadratic ally Constrained Quadratic Program

SISR	Single Image Superresolution
SR	Superresolution
SSIM	Structural Similarity Index Measurement

Chapter 1

INTRODUCTION

1.1 Introduction

Single image super-resolution is a type of inverse problem that is ill-posed. Many authors in literature have tried to regularize this problem by enforcing the constraints on the priors for the restriction of the solution space such as [30], [31], [38]. In most recent approaches sparsity has been effectively used as a prior, and by using this method many have achieved state-of-the-art results both visually and quantitatively. The sparse representation of a signal x , over a dictionary D have been clearly proven to be unique and reliable [28], [37], [26].

The signal-fitting property of sparse representation over learned dictionaries has been used extensively in sparsity based super-resolution approaches. This allows for reduced representation error. In the dictionary learning process one has a single feature space and the task is to design an over-complete dictionary. Yang et al. first gave the coupled dictionary learning model in [21] and improved it slightly in [11]. In this approach coupled dictionary training task is modeled as multi-variable optimization problem. In [21] a joint dictionary training method is proposed to train LR and HR patches by concatenating them to a single feature space. The dictionaries learned by this method are not actually for each feature space separately but they represent the concatenated feature space therefore during the testing phase reliable recovery is not guaranteed. The algorithm in [11] overcomes this problem in the

dictionary update by alternate optimization of the LR and HR dictionaries. Similarly in [10] coupled K-SVD algorithm is presented that further improves the coupling by using the best low-rank approximation given by the SVD. To get dictionaries with better representation power and improved coupling between LR and HR coefficients many authors have tried to improve the results of Yang's approach [40], [43], [42].

1.2 Motivation

The idea of designing multiple dictionaries instead of single one has already been proven useful for sparse representation of signals. In [44] Elad et al. has proposed a method to learn the multi-scale dictionaries using wavelets which helps better capture the intrinsic image features. Further in [41] the author has proposed the wavelet domain dictionary learning for single image super-resolution using K-SVD. It is noted that the directionality of wavelet subbands and the persistence of wavelet coefficients across scale [45], [46] are important attributes that can be utilized for dictionary learning and sparse representation. In the SISR framework the persistence property of wavelet coefficients [45] implies that similarity of sparse representation coefficients can be better exploited. However basic wavelet domain K-SVD based SR algorithm [41] cannot well take advantage of this property since it first learns the LR dictionary and then calculates the HR dictionary using the LR sparse representation coefficients and the HR training data via the pseudo inverse. The use of coupled K-SVD in the wavelet domain exploits the persistence property of the wavelet coefficients across different resolution levels, due to the dependency of the wavelet coefficients across different resolution levels using coupled K-SVD one can further enforce the similarity on the HR and LR subband sparse coefficients.

1.3 Thesis Contributions

In this thesis a dictionary learning algorithm based on coupled K-SVD [10] in wavelet domain is proposed and it is applied to the problem of Single Image Super-Resolution (SISR). Given the training image data set HR and LR patches are extracted for dictionary training. First a 2 level wavelet decomposition of the training images is performed to get the desired HR and LR wavelet coefficient subbands level 1 subbands are the HR subbands and level 2 are the LR subbands. For each of the horizontal, vertical and diagonal wavelet subbands, LR and HR dictionary pairs are designed using the coupled K-SVD algorithm. By using coupled K-SVD algorithm one can establish better coupling between the LR and HR feature spaces. This is achieved by enforcing that LR and HR sparse representations of the training data are similar. In coupled K-SVD this is done by updating dictionary pairs and corresponding sparse representation coefficients simultaneously. Given a set of HR and LR subbands for training, the LR subbands are interpolated such that they are the same size as the HR training images. Training patch pairs are prepared by extracting patches from the same spatial locations of the three wavelet subbands at two resolution levels. The patch pairs are then vectorized to form the training data for each wavelet subband. A dictionary size of 256 atoms and patch size of 6×6 with 20 coupled K-SVD iterations are used to learn the coupled dictionaries. For reconstructing the HR image, we assume that the given LR image is the same as the low resolution wavelet subband image. In other words we assume that the blur filter is the scaling filter of the wavelet transform. In this setting the reconstruction of the HR image amounts to estimating the horizontal, vertical and diagonal subbands and performing 1 level inverse wavelet transform.

Thus given the LR image a 1 level wavelet transform is performed to obtain the LR wavelet subbands from which the HR wavelet subbands are estimated via sparse representation. For each wavelet subband, first the sparse representation coefficients of the LR patches are calculated using the LR dictionary, then the corresponding HR patch is reconstructed by using the HR dictionary and the calculated sparse representation coefficients. The proposed algorithm produces better results when compared to the leading super-resolution algorithms. Its performance is quantitatively evaluated, and shown to possess a Peak-Signal to-Noise Ratio (PSNR) raise of 1.19 dB over the algorithm of [41], as run over the Kodak set and benchmark images. The average improvement over [11] is 2.24 dB. Furthermore the average improvement over the spatial domain coupled K-SVD algorithm [10] is 2.41 dB. This improvement is due to the coupled K-SVD dictionary training of the LR and HR patches in wavelet domain which results in better approximation of the HR subbands in overall reconstruction.

In this thesis we have developed a wavelet domain based Single Image Super-resolution (SISR) algorithm using the coupled K-SVD dictionary learning. The use of wavelet for the dictionary learning combined with the coupled dictionary learning produces good results when compared with the state of the art algorithms [10], [11], [41]. The use of coupled K-SVD for dictionary learning in the wavelet domain exploits a very important property of the wavelets which is persistence across scale of the wavelet coefficients. Taking advantage of this property of the wavelets the dictionaries are designed by enforcing the sparse representation coefficients of the two feature spaces in the super-resolution problem to be same. Using this approach the weak link between sparse representation coefficients is strengthened and in the

reconstruction process the trained dictionaries produce better results both quantitatively and qualitatively.

1.4 Thesis Outline

The remainder of the thesis follows as in Chapter 2 the super-resolution problem is discussed in more detail with the solution of joint sparse coding and coupled dictionary learning. In chapter 3 the proposed algorithm is presented with all the necessary details. Chapter 4 includes the simulations and results, where the proposed algorithm is compared with the state of the art algorithms and comparisons are shown quantitatively and qualitatively. Lastly in Chapter 5 the thesis is concluded and some future work is discussed. After that comes references section.

Chapter 2

LITERATURE REVIEW

2.1 Sparse Model

The method of sparse representation over learned dictionaries has become more popular in recent years and has become the hot topic of research for the researchers of signal and image processing. Many new algorithms and techniques have been developed based on this approach which solves a variety of problem including inpainting, de-noising, super-resolution, compression.

The sparse representation problem is divided into two steps: in the first step there is a basis which is known as dictionary and the second step sparse coding algorithm which is used to point out those basis which are then called atoms and the sparse coefficients, jointly they make the approximation of the signal. Since the sparse coding algorithm uses only a fraction of the atoms from the dictionary to represent the input signal hence it is named the sparse representation. For the application of this sparse representation there is a requirement on our input signal to be compressible. Compressibility is the representation of the input signal as a product of basis dictionary and the sparse coefficients.

2.1.1 Basic Idea

For a given image patch $x \in R^n$ of size $m \times m$ according to the Sparseland model [1] a matrix of size $D \in R^{n \times k}$ with $(k \gg n)$ which should be redundant. Now every x can be represented sparsely by this matrix D which is called the dictionary as,

$$\arg \min_{\alpha} \|\alpha\|_0 \text{ subject to } D\alpha \approx x \quad (2.1)$$

Where α is the sparse coefficient vector. Writing more clearly by changing the constraint with a clear bounded representation as error, the sparse representation now can be found by

$$\arg \min_{\alpha} \|D\alpha - x\|_2^2 + \mu \|\alpha\|_0 \quad (2.2)$$

with penalty

$$\arg \min_{\alpha} \|\alpha\|_0 \text{ subject to } \|D\alpha - x\|_2 \leq \epsilon \quad (2.3)$$

Where $\|\cdot\|_0$ and $\|\cdot\|_2$ represent the zero and two norm and ϵ represents the desired error.

Various attempts are made in literature to solve this NP-hard problem and many algorithms were developed such as Matching Pursuit and Basis Pursuit Algorithms and Relaxation method like L-1 norm minimization. Recent work suggests that these approximation techniques can be quite accurate if the solution is sparse enough to begin with. There are many other methods for the solution of this problem some of which are discussed here.

2.2 Orthogonal Matching Pursuit (OMP)

The OMP [2] is used to approximate the solution i.e. $\min_{\alpha} \|\alpha\|_0$. Here we have a matrix D and a vector y and a threshold ϵ . This is an iterative algorithm and its steps are; First we initialize the loop and the initial conditions as $\alpha^0 = 0$ as the initial solution $r^0 = x - D\alpha^0$ as the initial residual and $T^0 = \text{Support}\{\alpha^0\}$ as the initial support. Then in the main loop for each iteration we select the best atom from D with maximum projection on the residue and we keep it. Next is the update step where we find a minimizer for the desired threshold and update the support accordingly. After the updating of support we update the provisional solution according to the updated

support i.e. we compute the α^k the minimizer of $\|D\alpha - x\|_2^2$ subject the updated support, here we also update the residual by $r^k = x - D\alpha^k$. Finally the stopping rule is set by the threshold if the updated threshold is less than the desired threshold then we stop otherwise repeat the algorithm until the desired conditions are met.

2.3 Iterative Reweighted Least Squares (IRLS)

The IRLS [3] algorithm is used as a relaxation algorithm to solve the vector selection problem iteratively, it is used in Focal Underdetermined System Solver (FOCUSS) algorithm [4] which is minimum norm algorithm, it solves the following optimization problem,

$$\frac{1}{2}\|D\alpha - x\|_2^2 + \mu\|\alpha\|_1 \quad (2.4)$$

This algorithm works iteratively but here the relaxation is applied to the one norm instead of the zero norm. This algorithm works iteratively so we first initialize the initial parameters i.e. $a_0 = 1$ and weighted matrix $A_0 = 1$ as our initial parameters. Next for each iteration we first find the approximate solution of the linear system $(2\mu A_{k-1}^{-1} + D^t D) = D^t x$ iteratively to produce the result a_k . Next we update the weight matrix by using the approximated value of a_k as $A_k(j, j) = |A_k(j)| + \epsilon$. Finally the stopping condition is set based on the error term $\|a_k - a_{k-1}\|_2$.

2.4 Least Angle Regression Stage-wise (LARS)

The LARS [5] algorithm was proposed by the LASSO [6] team at Stanford, this algorithm is also a relaxation algorithm which is quite similar to the OMP algorithm. Here we solve the same L1- Norm minimization problem given by

$$\frac{1}{2}\|D\alpha - x\|_2^2 + \mu\|\alpha\|_1 \quad (2.5)$$

This algorithm works on a sub-gradient method to solve this minimization problem

$$\partial f(a) = \{D^t(Da - x) + \mu z\} \forall z = \begin{cases} +1, & a[i] > 0 \\ [-1, +1], & a[i] = 0 \\ -1, & a[i] < 0 \end{cases} \quad (2.6)$$

The main goal of this algorithm is to seek the value of a and z so that 0 belongs to this sub-gradient. The main core of this algorithm is described below,

1. First the sparsity regularizer is set $\mu > \infty$ and we assume $a_\mu = 0$ as our optimal solution and we can observe as μ approaches zero then our zero solution remains optimal. Using this and that zero should be in the sub-gradient we say that $0 = -D^t x + \mu z_\mu$ where $z_\mu = \frac{D^t x}{\mu}$.

2. As μ is decreasing and getting to $\|D^t x\|_\infty$, avoiding to violate the sub-gradient condition and keeping the optimal solution intact we set $z_\mu = \text{sign}[a_\mu\{i\}]$ and we can get it by

$$0 = a_i^t(aA - b) + \mu z_\mu[i] \quad (2.7)$$

$$a_\mu[i] = \frac{a_i^t y - \mu z_\mu[i]}{a_i^t a_i}$$

$$a_\mu[i] = \frac{a_i^t y - \mu \text{sign}[a_\mu\{i\}]}{a_i^t a_i}$$

Now this solution is valid for sparsity regularizer and all values near to it.

3. Next the solution goes smoothly and linearly by decreasing the value of sparsity regularizer and updating z_μ as $z_\mu = \frac{D^t(y - Da_k)}{\mu}$ with some support of a setting μ fixed given by S , a_μ^S is existing part of a_μ as non-zero. With the matrix set as to represent the data chosen called the sub-gradient matrix, the solution is given by,

$$a_\mu^S = (D_S^t D_S) (D_S^t x - \mu z_\mu^S) \quad (2.8)$$

Where $z_\mu^S = \text{sign}(a_\mu^S)$, the sparsity regularizer is decreases step by step with calculating these equations togher with updating the support and solution a_μ and z_μ .

2.5 Dictionary Learning

In the sparse representation we consider our signal of interests are assumed to be compressed. This is the basic idea when we represent the signals of interest sparsely. Compressibility is basically the use of the over-complete dictionaries also called the bases in some cases to represent the data, where a very small number of dictionary atoms are used in the data representation. In signal compression the main purpose is to represent the signal class using the specific dictionaries of interest, and in the literature various authors have attempted to design algorithms to train these dictionaries on a particular signal class of interest. The advantage of using the trained dictionaries instead of the fixed ones becomes evident when the trained dictionaries used give the low approximation error when compared with the fixed ones. The training signals can be represented sparsely as;

$$E = \|X - DA\|_F^2 \quad (2.9)$$

Let $\|\cdot\|_F$ denote the Frobenius norm and X is the data matrix and A is the sparse representation matrix. The coefficient matrix is calculated given the dictionary and the data by a vector selection algorithm such as OMP, LARS. The dictionary training methods are iterative and switch between two stages one is the vector selection problem which is more expensive and other is the dictionary update stage and for each iteration both these quantities are evaluated by keeping the other fixed. The main difference between these algorithms is the steps to find the sparse coefficients and the method for updating the dictionary.

Olshausen and Field proposed a dictionary training algorithms [7]. In this method, maximum-likelihood estimation is used to find the optimal dictionary. They used the Gaussian or Laplace method for the sparse coefficient approximation and used the

concept of steepest descent to update the dictionary and the sparse coefficients matrix.

2.5.1 Method of Optimal Direction Algorithm (MOD)

This algorithm works iteratively alternating between two stages one is the sparse coding and other is the dictionary update stage. A dictionary D is trained on the data $\{y_i\}$ by approximating the sparse representation problem as described in the previous section. The algorithm is described in the following steps,

1. First the initialization is carried by setting the initial parameters of the algorithm, a dictionary D is initialized randomly as $D_0 \in R^{n \times m}$ either by using random entries or by choosing the random samples from data, next it is normalized.
2. Next for each iteration sparse coding is applied first by using one of the sparse coding algorithms described previously i.e.

$$\arg \min_a \|x_i - D_{k-1}a\|_2^2 \quad \text{subject to } \|a\|_0 \leq T_0 \quad (2.10)$$

This constitutes A_k matrix.

3. After the sparse representation coefficients have been approximated using an initial guess dictionary and the given data, the next step is the dictionary update stage. In the MOD algorithm the dictionary atoms are updated by a simple formula given as

$$\arg \min_A \|X - DA_k\|_F^2 = XA_k^t (A_k A_k^t)^{-1} \quad (2.11)$$

Finally the stopping criteria is set based on the error term $\|X - D_k A_k\|_F^2$, if the error term is small enough we stop the algorithm otherwise iterate until convergence. At the end the desired output is obtained.

2.5.2 K-SVD Algorithm

This algorithm was given by Aharon et al [9], here again the task is same we have to train a dictionary based on the data and sparse represent the data by approximating

the solution of the NP-hard problem. The algorithm works in similar way iteratively we chose the initial dictionary, normalize it and for each iteration we first find the sparse solution of the optimization problem using OMP algorithm which gives the sparse representation coefficients. Using the sparse representation coefficients, we have to update the dictionary atoms in the next step. For that first we find the group of samples from the data that use the specific atoms in the dictionary as;

$$\Omega_{j_0} = \{i | 1 \leq i \leq M, A_k[j_0, i] \neq 0\} \quad (2.12)$$

Then we compute the residual matrix as;

$$E_{j_0} = X - \sum_{j \neq j_0} d_j a_j^t \quad (2.13)$$

The next step is to restrict the residual matrix according the index matrix created to track the samples using the atoms of the dictionary. After that apply SVD decomposition on the restricted residual matrix as $E_{j_0}^R = U\Delta V^T$, then set update of atom and sparse representation as $d_{j_0} = u_1$ and $a_{j_0}^R = \Delta[1,1].v_1$. u_1 represents the first column of the orthogonal matrix U from SVD and v_1 represents the first column of the orthogonal matrix V from the SVD. We stop the algorithm by detecting the change in the error term $\|X - D_k A_k\|_F^2$ if it is negligible we stop otherwise perform another iteration. We keep iterating until we get the desired output.

2.6 Applications

There are a many applications which use this approach of dictionary learning using sparse representations; we describe a few of them here.

2.6.1 Inpainting

The use of inpainting in signal processing is extensive and wide. In this application we estimate the missing pixels from a given image. Its use is mostly in the data transmission where for the channel codes a replacement is provided [12], [13] and

also it is used to remove the superimposed text in the image manipulation, signboards and highways or marketing banners [14].

Suppose we have a patch from the image defined as $[x^T x_k^T]^T$ in which x is the available data and x_k is the missing data which will be estimated using the image inpainting technique. In [15] author has proposed a method of estimating it, in which a concatenation of orthonormal basis is employed by some compressibility method. In compressibility a vector can be estimated from a dictionary and some sparse representation vector. The method to recover the missing data is defined by the following equation where they set the diagonal matrices Δ_n , Δ_a , Δ_m which define the non-zero entries of the sparse representation matrix, the existing data in the image and last one is the missing data from the image as follows;

$$\hat{x}^i = \begin{pmatrix} x_a \\ \hat{x}_m^i \end{pmatrix} = \Delta_a \hat{x}^{i-1} + \Delta_m (D \Delta_n D^T) \hat{x}^{i-1} \quad (2.14)$$

2.6.2 Denoising

One of the most useful applications of sparse representation is denoising on images and videos [15], [16]. Considering a sparse prior on data the image denoising is basically the MAP estimation of data. In the MAP estimation the sparse representation of images is probably in blocks and overlapped is computed, then the MAP solution is estimated and the result is given by averaging of the all blocks of images.

The problem of image denoising using sparse representation is very simple; given a noisy image we first extract patches from the image in an overlapping manner then applying dictionary learning on it using K-SVD or any other method and finding the sparsely coded matrix by OMP or some other vector selection method. In this

process the best dictionary atom represents the useful part of the image and the noisy parts gets rejected, lastly the patches are converted to the $2 - D$ patches by overlapping and the reshape and merge operation is applied to get the denoised image.

2.6.3 Image Compression

Sparse representation for image compression can be employed more efficiently because of the process of learned dictionaries which can be used to represent a signal class very efficiently. An example of this technique can be considered as the work done by [17] based on the dictionary learning. Here they learn the dictionaries on the class of face templates which are non-overlapping using the K-SVD approach and the results are much improved in comparison with the existing JPEG2000[18] algorithm with the added disadvantage of increase in the storage size of the data.

2.6.4 Image Superresolution

Image superresolution is a very important and active area of research because of its demand in many applications. To recover the loss of high frequency information from an image encountered during data acquisition, transmission or storage is known as superresolution. Generally superresolution is employed by using several low resolution images however if the available images are limited then these methods become impractical, here comes the concept of Single Image Superresolution (SISR).

Previously many authors have attempted to solve this problem, but the state of the art became Yang et al [19] and Zeyde et al [20]. In the Yang's method of SISR [19], [21] the image patches are extracted from a pre-defined data set of images and classified as Low Resolution (LR) and High Resolution (HR) images, then separate dictionaries are learned on these images. The superresolution is achieved by

assuming the sparse coefficients of LR and HR patches to be the same and the HR patches are recovered by applying the HR dictionaries on the sparse coefficients of the LR patches.

Chapter 3

SUPER-RESOLUTION

3.1 Introduction

In most basic case we want a high resolution image from a low resolution image. The low resolution image is most probably the one which have lost its high energy components during some processing. In the super-resolution approach two main constraints are observed, one is the reconstruction constraint and other is the sparsity prior. In the reconstruction constraint the reconstructed HR image must be in agreement with the LR image based on the observation model and in the sparsity prior constraint the HR image which is easily represented sparsely over dictionary and it can be reconstructed from the LR image.

Consider a LR image Y which we get by down sampling and blurring a HR image X , now we assume that the HR patches can be represented sparsely by an over-complete dictionary as $x = D_h \alpha_0$ where $\alpha_0 \in R^k$ is the sparse representation vector which has very less number of non-zeros ($\ll k$). Now we can develop a relation between a HR and LR patch as,

$$Y = SBx = Lx = LD_h \alpha_0 \quad (3.1)$$

Where S is the down sampling operator, B is the blurring operator and L is their combined effect. Assuming that this $D_l = LD_h$ we can get

$$y = LD_h \alpha_0 = D_l \alpha_0 \quad (3.2)$$

The above equation suggests the similarity between the LR and HR sparse representation coefficients. Now the HR patch reconstruction becomes very simple. Given the LR patch we can obtain its sparse coefficients and using the similarity principle we can get the HR patch as

$$\hat{x} = D_h \hat{\alpha}_0 \quad (3.3)$$

The sparse representation is the vector selection problem which is NP-Hard and not easily solvable and it is computationally expensive. It is an optimization problem and can be given as;

$$\min_{\alpha_0} \|y - D_l \alpha_0\|_2 \text{ s.t. } \|\alpha_0\|_0 < T_0 \quad (3.4)$$

Where T_0 is the sparsity parameter and $\|\cdot\|_2$ and $\|\cdot\|_0$ represent the two norm and zero norm.

For the representation of our signal we need a suitable dictionary and a sparse matrix. This problem basically a two-step problem, in the first step we represent the sparse coefficient matrix from an initial guessed dictionary which can be achieved by any vector selection algorithm, the other step is the update of the dictionary using the sparse representation and given data which can be done by any dictionary learning algorithm.

3.2 Joint Sparse Coding

This involves learning the two coupled dictionaries for the coupled feature spaces which Yang et al [11] described as latent space and the observation space. He also proposed a method to solve this as;

$$\min_{D_x D_y [a_i]} \sum_{i=1}^N \|x_i - D^x a_i\|_2^2 + \|y_i - D^y a_i\|_2^2 \quad (3.5)$$

$$\text{s.t. } \|a_i\|_0 \leq T_0, \|d_r^x\|_2 \leq 1, \|d_r^y\|_2 \leq 1$$

$$i = 1, 2, \dots, N, r = 1, 2, \dots, n$$

This formulation suggests that the obtained sparse representation should be able to recover both x_i and y_i very well. What is done here that is the latent and observation spaces are joined together to form a concatenated space and the error terms are represented as;

$$\hat{x}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$$

$$\hat{D} = \begin{pmatrix} D_x \\ D_y \end{pmatrix}$$

Now this becomes a standard sparse coding problem and can be represented as;

$$\min_{\hat{D}, \{\alpha_i\}_{i=1}^N} \sum_{i=1}^N \|\hat{x}_i - \hat{D}\alpha_i\|_2^2 + \mu \|\alpha_i\|_1 \quad (3.6)$$

$$s. t \ \|\hat{D}(:, k)\|_2 \leq 1$$

The above optimization problem more looks like an optimization for the joint feature space not the single feature space, therefore one cannot assume such solution to be optimal for each feature space separately.

3.3 Coupled Dictionary Learning

As proposed by the Yang et al [11] consider two feature spaces one is called latent and other is the observation space and both are coupled together and denoted as $X \in R^{d1}$ and $Y \in R^{d2}$. The signals are assumed to be sparse over certain sparse representations over the learned dictionaries. Thus one can infer the elements of X for the elements of Y which are the observations.

There dwells a mapping function between these two spaces which is not known and it is not linear, which maps a signal element from the latent space to the observation space as; $F: X \rightarrow Y$ and $Y = F(x)$, for the inference to be possible from y to x this mapping function is assumed to be nearly injective. The main problem here is to find a pair of dictionary for each space i.e. Latent and Observation space such that we can

estimate the signal in the latent space from the observation space. The signal in the latent space is represented in terms of the dictionary D_x and in the observation space it can be represented in terms of the dictionary D_y .

For a coupled signal $[y_i, x_i]$ the learned joint dictionaries D_x and D_y must obey;

$$z_i = \arg \min_{\alpha_i} \|y_i - D_y \alpha_i\|_2^2 + \mu \|\alpha_i\|_1 \quad \forall i = 1, 2, \dots, N \quad (3.7)$$

$$z_i = \arg \min_{\alpha_i} \|x_i - D_x \alpha_i\|_2^2 \quad \forall i = 1, 2, \dots, N$$

Here we impose the similarity constraint of sparse coefficients of the two spaces where $\{x_i\}_{i=1}^N$ are the samples from the latent space and $\{y_i\}_{i=1}^N$ are the samples from the observation space and $\{z_i\}_{i=1}^N$ are the sparse representation matrices. This problem is similar to the compressed sensing problem [22]. In compressed sensing case the mapping is a linear random projection function. The dictionaries chosen for the D_x are the defined basis such as wavelets and we compute the D_y directly from D_x with this linear mapping. However one cannot apply the compressed sensing theory if the linearity of the mapping functions is unknown, therefore as an alternate solution one can use machine learning to train the data to obtain the dictionaries in coupled manner.

In this joint dictionary process the goal is to find the latent signal x from the observation y , here the main motive of this process is to minimize the error of recovery of estimated signal,

$$L(D_x, D_y, x, y) = \frac{1}{2} \|D_x z - x\|_2^2 \quad (3.8)$$

Hence we minimize this as

$$\min_{D_x, D_y} \frac{1}{N} \sum_{i=1}^N L(D_x, D_y, x_i, y_i)$$

$$s.t \ z_i = arg \min_{\alpha} \|y_i - D_y \alpha\|_2^2 + \mu \|\alpha\|_1 \quad \forall i = 1, 2, \dots, N \quad (3.9)$$

$$\|D_x(:, k)\|_2 \leq 1, \|D_y(:, k)\|_2 \leq 1 \quad \forall k = 1, 2, \dots, K$$

Now this is a highly nonlinear and non-convex problem so we minimize it alternately keeping one fixed and minimizing the other. The minimization of D_x is given as,

$$\min_{D_x} \sum_{i=1}^N \frac{1}{2} \|D_x z_i - x_i\|_2^2$$

$$s.t \ z_i = arg \min_{\alpha} \|y_i - D_y \alpha\|_2^2 + \mu \|\alpha\|_1 \quad \forall i = 1, 2, \dots, N \quad (3.10)$$

$$\|D_x(:, k)\|_2 \leq 1 \quad \forall k = 1, 2, \dots, K$$

One can solve this by the conjugate gradient descent [23]. Now minimizing D_y keeping the D_x fixed is a highly non-convex bi-level programming problem [24]. It can be minimized by the gradient descent method developed in the [25].

In this gradient descent method a stochastic gradient procedure is employed for the optimization of the D_y . Because of the fact that this optimization is highly non-convex bi-level, this method will only find a local minimum of the solution which is sufficient in practice.

3.4 The Coupled K-SVD Algorithm

The coupled K-SVD algorithm proposed by Xu et al [10] modifies the dictionary training given by the Yang et al [11] which is as follows;

$$\min_{D_x D_y [a_i]} \sum_{i=1}^N \|x_i - D^x a_i\|_2^2 + \|y_i - D^y a_i\|_2^2$$

$$s.t \ \|a_i\|_0 \leq T_0, \|d_r^x\|_2 \leq 1, \|d_r^y\|_2 \leq 1 \quad (3.11)$$

$$i = 1, 2, \dots, N, r = 1, 2, \dots, n$$

where d_r is the r^{th} atom of the dictionary and r_i is the sparse representation matrix, in the coupled K-SVD dictionary learning algorithm one first proposes a pair of dictionaries by randomly selecting the patches from the data and start the iteration. The sparse representation coefficients are approximated at the beginning of iterations for the both the dictionaries separately and alternately. Then the calculated sparse coefficients are used as the initial common sparse representation for the dictionary update stage. Next each atom of the coupled dictionary pair is updated.

In the dictionary update stage first a common sparse representation matrix is set for the coupled data, then according to the K-SVD algorithm the samples from the data that use the corresponding atoms of the dictionaries are located, next the coupled data according to that indices matrix containing location information of atoms is restricted. Now the residual matrices for the coupled data are set as done in the K-SVD and update the dictionary and the representation by applying the SVD to the residual matrices for each iteration. In this algorithm, at the beginning of each iteration the sparse representation matrix is calculated for the high resolution data and low resolution data alternately so that their individual features can be exploited in the overall learning process.

Chapter 4

THE PROPOSED SUPER-RESOLUTION METHOD

4.1 Single Image Super-Resolution

The main idea in sparse representation based SISR is that a HR patch x_H and a LR patch x_L can be sparsely coded over a HR dictionary D_H and a LR dictionary D_L as

$$x_H \approx D_H \alpha_H \quad (4.1)$$

$$x_L \approx D_L \alpha_L \quad (4.2)$$

where α_H and α_L are the representation coefficient vectors of x_H and x_L respectively.

The relationship between the HR and LR images can be characterized by the blurring and down-sampling operator ψ , as

$$x_L = \psi x_H \approx \psi D_H \alpha_H \approx D_L \alpha_L \quad (4.3)$$

Incorporating this assumption in (1) and comparing to (3), yields

$$x_H \approx D_H \alpha_H \approx D_H \alpha_L \quad (4.4)$$

This means that the sparse representation of the HR patch over its dictionary is approximately equal to that of corresponding LR patch over a LR dictionary. The approaches used in [11] and [10] design two dictionaries by coupled dictionary training of LR and HR patches. Then, each patch of the LR image is sparsely coded over the LR dictionary. The HR patch reconstruction is thus straight forward; each HR patch is reconstructed by the sparse representation coefficients of the corresponding LR patch, and the HR dictionary, As

$$x_H \approx D_H \alpha_H \approx D_H \alpha_L \quad (4.5)$$

In the prior work by [36] the super-resolution problem was approached by learning a separate LR dictionary for each LR subband in the wavelet domain using [9] and assuming the sparse coefficients of LR and HR patches to be the same and estimating the HR dictionaries by the pseudo inverse of HR patches and the LR sparse coefficients. This approach serves a weak coupling between the LR and HR coefficients is weak. This can be overcome by coupled dictionary learning of the LR and HR patches as in [11] and [10]. In this work K-SVD based coupled dictionary training [10] in the wavelet domain is employed.

4.2 Proposed SR Approach

Designing of multiple dictionaries instead of one has been proved to be effective by [29]. Moreover, use of clustering and classification [27], [34] to categorize the training data into several sets has been used effectively. These data sets are then used for designing of the dictionaries. Discrete Wavelet Transform (DWT) can split image information into horizontal, vertical and diagonal features. Coupled K-SVD algorithm is used in the wavelet domain for SISR. The coupled K-SVD algorithm achieves better coupling of the LR and HR sparse coefficients by enforcing the LR and HR dictionary atom pairs to have the similar indices also it uses a single sparse representation to update the LR and HR dictionary pair atoms. It does this in an iterative manner by alternatively selecting the LR and HR sparse coefficients and using them for the single sparse representation of both. The coupled K-SVD better exploits the persistence property of the wavelet coefficients at different resolution levels. A strong dependency is present in the wavelet coefficients across different levels. The use of coupled K-SVD for training on the wavelet subband images at the 2 levels produces better coupling in their sparse representation coefficients.

4.2.1 The Proposed Dictionary Learning Algorithm

For the training of the coupled dictionaries, we first extract a large number of HR/LR image patch pairs from some predefined database containing clean images. For each HR image two-level DWT decomposition is performed to generate the training detail subbands. These are referred to as I_L^y and I_H^y where the subscript H(L) stands for HR(LR) and the subscript $y = \{h, v, d\}$ refers to the specific wavelet subband. h, v, d represent horizontal, vertical and diagonal subbands respectively. The level-one (level-two) subbands are assumed to be the HR (LR) training data. LR subbands are then interpolated by wavelet interpolation to increase their size and match with the size of the HR image as done in many approaches [20], [11] and [10]. This image is referred to as the Mid Resolution (MR) image and is denoted by I_M^y .

For each subband denoted by $y = \{h, v, d\}$, we sample pairs of HR/LR patches, and vectorize to form the training data matrices W_H^y and W_M^y . It is pointed out that in our approach there is no need to use the feature extraction filters. Patches with small variances are eliminated. Then the LR and HR dictionaries D_L^y and D_H^y are jointly trained with the MR and HR subband training data matrices W_H^y and W_M^y . The dictionaries and the sparse representation coefficients are obtained by solving the following optimization problem.

$$\begin{aligned} \min_{D_L^y, D_H^y, \alpha(i)} \sum_{i=1}^N & \|W_M^y(i) - D_L^y \alpha(i)\|_2^2 + \|W_H^y(i) - D_H^y \alpha(i)\|_2^2 \\ \text{s.t. } & \|\alpha(i)\|_0 \leq T_0, \|D_L^y(k)\|_2 \leq 1, \|D_H^y(k)\|_2 \leq 1 \\ & i = 1, 2, \dots, N, \quad k = 1, 2, \dots, K, \end{aligned} \quad (4.6)$$

where K is the dictionary size, $D_H^y(k)$ and $D_L^y(k)$ are the k^{th} atom and $W_M^y(i)$ and $W_H^y(i)$ are the i^{th} training vector of the HR and LR dictionaries respectively and T_0

is the level of sparsity. The coupled K-SVD algorithm [10] is used to learn the dictionaries.

The horizontal and vertical and diagonal subbands are directly fed to the coupled K-SVD algorithm to generate the LR and HR horizontal, vertical and diagonal subband dictionaries. This process is summarized in Figure 4.2.

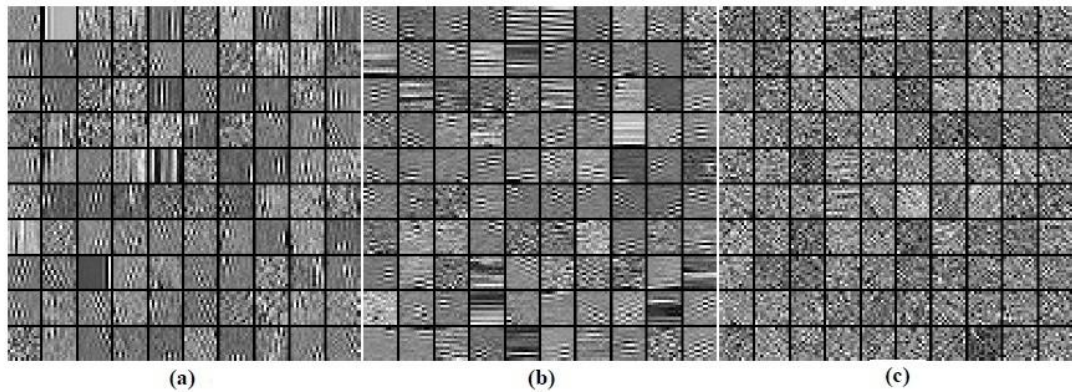


Figure 4.1: HR Dictionaries learned using coupled K-SVD
(a) Diagonal Structured (b) Horizontal Structured (c) Vertical Structured

Figure 4.1 shows some atoms from the designed HR subband dictionaries. It can be clearly observed that the designed dictionaries exhibit the horizontal, vertical and diagonal structure as they were trained on the corresponding horizontal, vertical and diagonal wavelet subbands from the training data.

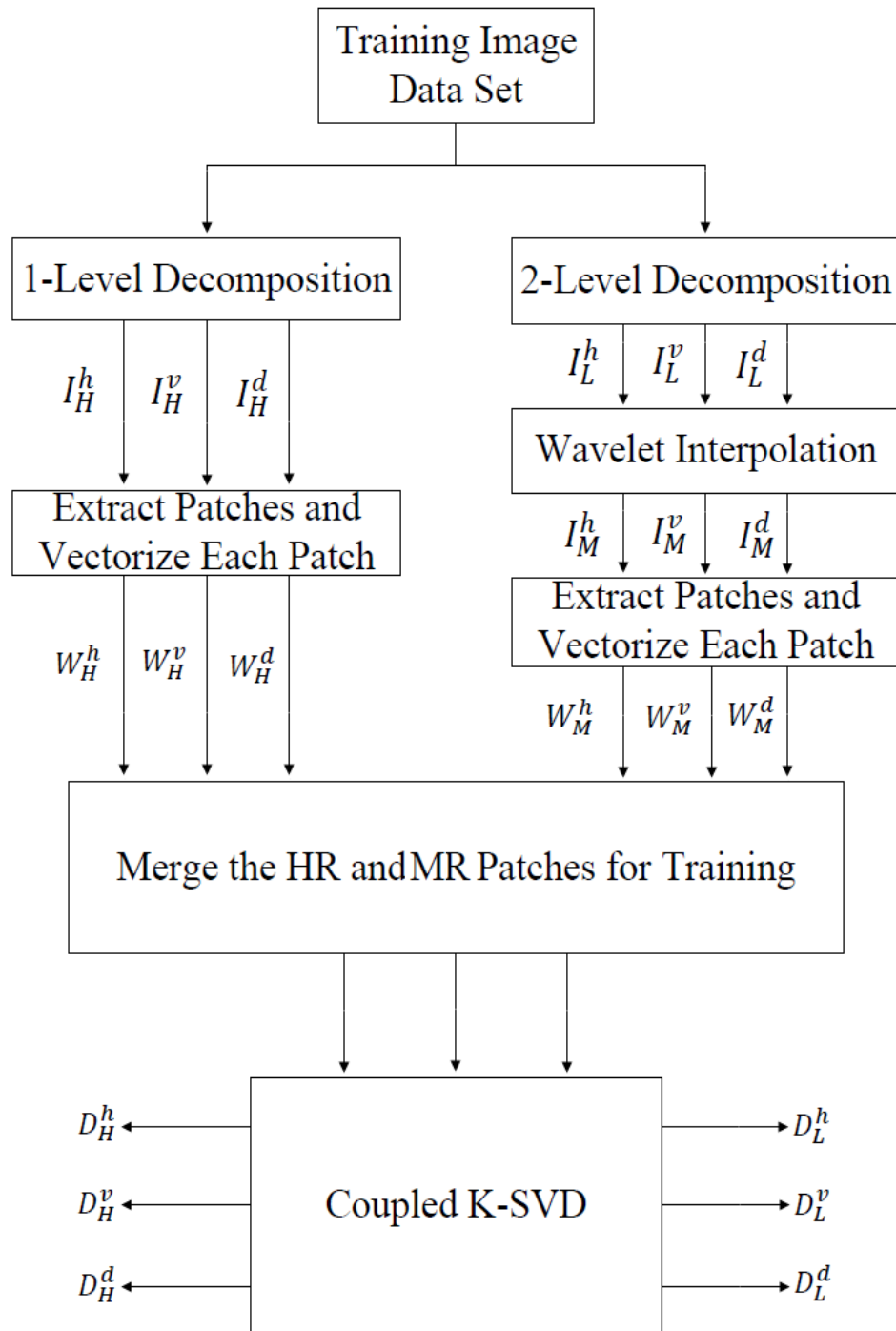


Figure 4.2: Proposed Dictionary Learning Algorithm

4.2.2 The Training Image Dataset

For the training of dictionaries in the wavelet domain we have used the training image data set given by the [21] which is also used by the Elad et al. [20] This set consists of a set of natural images rich in high frequency contents and are small in dimensions. This standard training set used by most of the researchers for the dictionary learning process. In our proposed algorithm for dictionary learning in the wavelet domain first we do the 2 level wavelet decomposition on these images and then we sample the patches for each of the DWT subband i.e. horizontal, vertical and diagonal for both the feature spaces HR and LR. Some of the images used for the training are shown in the Figure 4.3.



Figure 4.3: Samples of images used in dictionary learning [21]

4.2.3 The Proposed Image Reconstruction Algorithm

The purpose of this algorithm is to reconstruct the original HR image by estimating its detail wavelet subbands. Here the LR image is assumed to be the low resolution (approximation) subband of the wavelet transform of the HR image. In essence this assumption is equivalent to assuming that the blurring operator is the low pass scaling filter of the wavelet transform, one can design an orthogonal or even a biorthogonal wavelet filter bank where the low pass scaling filter approximately matches the blurring operator. Here we used symlets [39] wavelets for DWT analysis and synthesis. In this setting one needs to estimate the wavelet subbands of the LR image and employ a 1 level inverse DWT to reconstruct the HR image. Thus in line with the dictionary learning algorithm a 1 level forward DWT of the LR image is taken. Then a wavelet based interpolation is performed to each wavelet subband to make them have the same size as the (unknown) wavelet subbands of the HR image to be reconstructed. The sparse representation coefficients of the patches extracted from each subband is calculated by solving the following optimization problem.

$$\arg \min_{a_L^y} \|W_M^y - D_L^y a_L^y\|_2 \text{ s.t. } \|a_L^y\|_0 < T_0 \quad (4.7)$$

This is a vector selection process and one can solve it using the greedy algorithms such as Orthogonal Matching Pursuit (OMP) [2]. Then using the calculated sparse representation coefficients and the corresponding HR dictionaries the wavelet subbands of the HR image are estimated as,

$$W_H^y = D_H^y a_H^y \approx D_L^y a_L^y \quad (4.8)$$

First reconstructed vector signals W_H^y are reconstituted as the $2 - D$ patches and then the merge method of [11] is used to constitute the full wavelet subband images. This process is summarized in Figure 4.4.

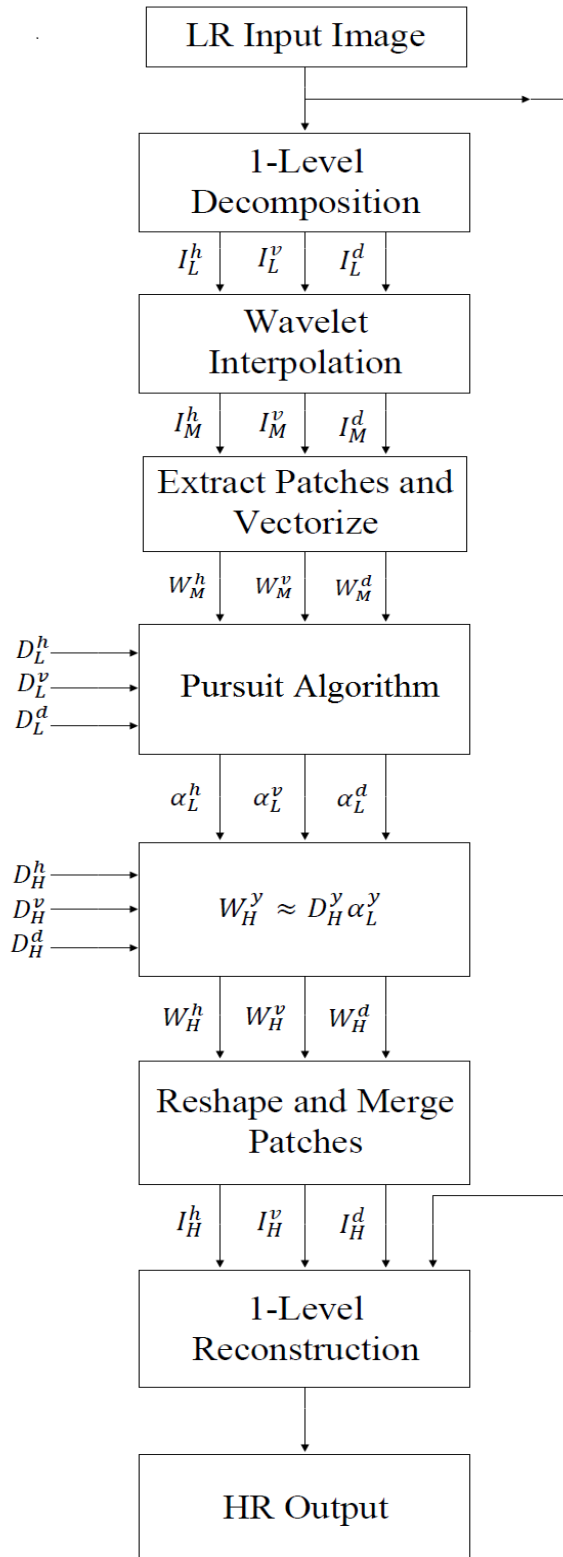


Figure 4.4: The Proposed Image Reconstruction Algorithm

Chapter 5

SIMULATION AND RESULTS

5.1 Introduction

The proposed wavelet domain coupled K-SVD algorithm is compared with three algorithms, the algorithm by Yang et al.[11], the algorithm by Xu et al.[10], the algorithm by Nazzal et al.[41] and the bicubic technique. Table 1 shows the PSNR and SSIM [46] results for the compared algorithms. The proposed algorithm uses a patch size of 6×6 with 256 dictionary atoms for each wavelet subband dictionary. The baseline algorithm of Yang et al.[11] and the algorithm by Xu et al.[10] use a patch size of 6×6 with 1000 dictionary atoms for single spatial domain dictionary. The algorithm by Nazzal et al.[41] uses a patch size of 6×6 with 216 dictionary atoms for each of the wavelet subband dictionary. There is a patch overlap of 5 pixels for all the algorithms. Images in all the compared algorithms are super-resolved with scale up parameter 2. Simulation is carried out for all the algorithms by setting all other parameters same to avoid any unfair advantage. For the implementation of the bicubic technique we have used Matlab's (imresize) function.

The proposed algorithm produces better results when compared with [11], due to the wavelet domain dictionary learning. Wavelet domain dictionaries exhibit properties of wavelets such as compactness and directionality and are small sized. The proposed algorithm gives an average PSNR raise of 2.24 dB over the state of the algorithm of [11] with SSIM improvement of 0.1006 tested on some Kodak data set and

benchmark images. This algorithm also outperforms the spatial domain coupled K-SVD algorithm of [10] with PSNR raise of 2.41 dB and SSIM improvement of 0.1028. This justifies the fact that wavelet domain dictionaries better recover some of the high frequency components of the LR image with low cost.

The comparison with the algorithm of Nazzal et al.[41] is made because it uses wavelet domain dictionary learning with K-SVD (no coupling). Simulation results show an average PSNR improvement of 1.19 dB and a slight increase in average SSIM of 0.0074. This shows that the coupling in dictionary learning can better exploit the persistence property [45] of the wavelets. The results are improved significantly due to better coupling of HR and LR sparse representation coefficients in wavelet domain by coupled K-SVD. The PSNR and SSIM improvement over the bicubic technique are 4.01 dB and 0.0986.

5.2 Quantitative Result

Table 5.1 lists the PSNR values of the Kodak data set and other benchmark images (gray level versions) reconstructed with bicubic interpolation,[11],[10] Algorithm, [41] and the Proposed algorithm, respectively. According to simulations, a patch size of 6×6 , and a dictionary size of 256 is used. Results are obtained by 20 K-SVD iterations. The training set includes natural and artificially-generated images. It is made sure that the image to be super-resolved is not included in the training set.

Table 5.1: The PSNR (up) and SSIM (down) value comparison of the proposed algorithm with the Bicubic, Yang et al, Xu et al, Nazzal et al algorithms

Images	Bicubic	Yang et al	Xu et al	Nazzal et al	Proposed
Kodak01	26.58	27.88	27.63	30.1	31.59
	0.7429	0.8283	0.8222	0.9726	0.9750
Kodak02	33.17	34.1	34.02	36.59	37.01
	0.8617	0.9039	0.902	0.9834	0.9864
Kodak03	33.95	35.37	35.1	36.92	37.56
	0.9115	0.9446	0.9419	0.9818	0.9914
Kodak05	26.98	29.09	28.67	30.31	32.07
	0.8343	0.9055	0.8973	0.9869	0.9899
Kodak09	32.61	34.6	34.18	35.07	36.28
	0.9028	0.9375	0.935	0.9836	0.9874
Kodak10	32.46	34.34	33.83	36.11	36.48
	0.8996	0.9387	0.9339	0.9901	0.9936
Kodak11	29.69	31.02	30.79	32.26	34.26
	0.826	0.8836	0.8791	0.9804	0.9819
Kodak12	33.08	34.76	34.52	37.09	38.23
	0.8812	0.9204	0.9183	0.9835	0.9876
Kodak14	29.47	30.95	30.72	33.26	33.68
	0.8324	0.8928	0.8893	0.9879	0.9889
Kodak18	28.72	29.96	29.84	31.23	33.63
	0.8325	0.8931	0.8911	0.9816	0.9850
Kodak23	34.72	36.39	36.25	39.04	39.34
	0.9467	0.9643	0.9638	0.9857	0.9959
Baboon	22.98	25.71	25.67	24.61	26.64
	0.933	0.7926	0.7912	0.9796	0.9798
Barbara	25.27	25.82	25.8	25.73	26.31
	0.9117	0.8446	0.8437	0.968	0.9799
Boat	29.93	33.95	33.74	33.76	34.99
	0.9276	0.9326	0.93	0.9741	0.9872
Elaine	31.06	31.33	31.33	31.45	33.97
	0.9088	0.7128	0.7135	0.9687	0.9906
Fingerprint	31.95	34.28	34.49	34.98	35.88
	0.9911	0.9717	0.9731	0.9974	0.9978
Lena	34.7	36.96	36.84	36.8	37.34
	0.9566	0.946	0.9456	0.9902	0.9946
Peppers	30.28	34.15	34.02	34.52	36.81
	0.9505	0.9179	0.9165	0.9831	0.9905
Zone Plate	11.4	12	12	12.72	13.12
	0.6923	0.5744	0.5755	0.7978	0.8375
Average	29.42	31.19	31.02	32.24	33.43
	0.8812	0.8792	0.877	0.9724	0.9798

5.3 Qualitative Result

Visual comparison is shown in figure 5.1 and 5.2. Reconstructed images obtained with the proposed,[11],[10],[41]and bicubic interpolation for the image number 1 in the Kodak set and Boat image. They are shown in sub-figures respectively. Also the accompanying insets compare an example comparative scene (best viewed on a computer screen). In line with the PSNR values, the proposed algorithm's reconstruction is improved than [10] and [11].

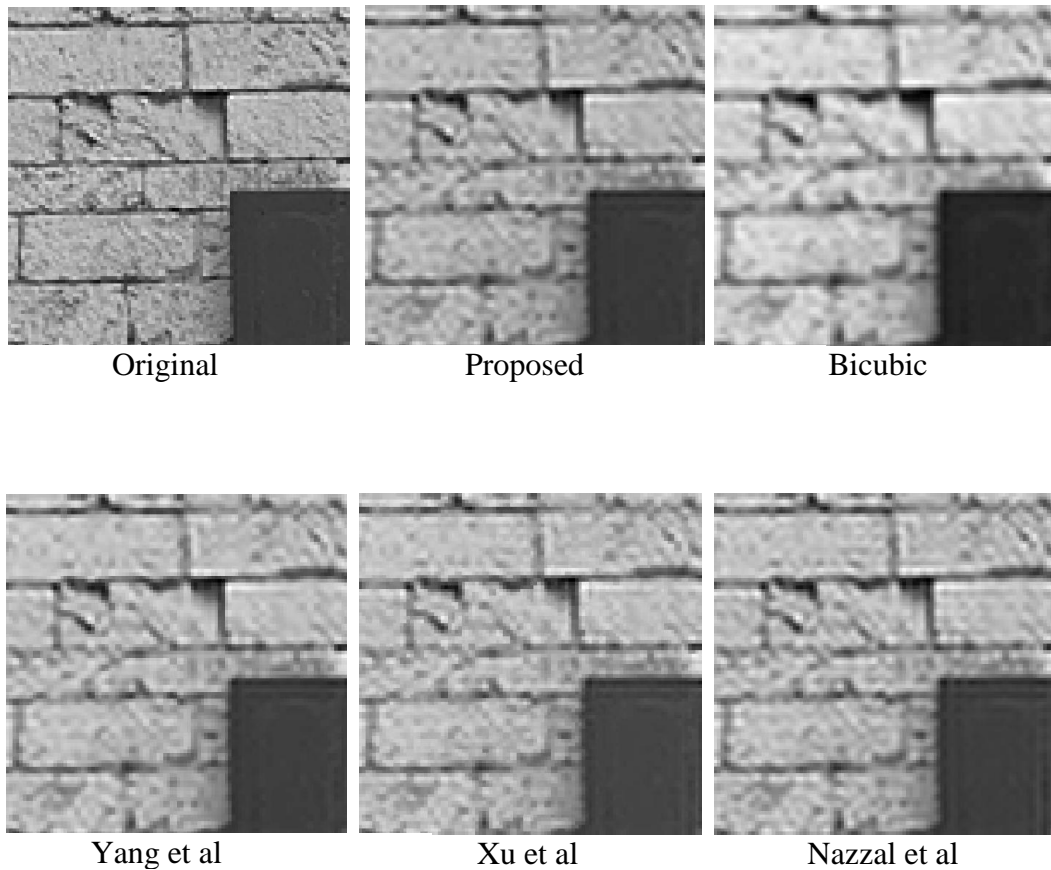


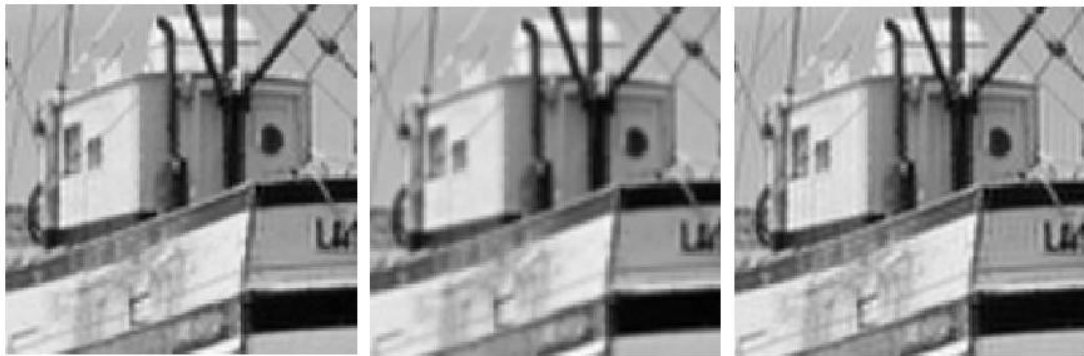
Figure 5.1: Visual Comparison of the Image 1 from the Kodak data set with the Bicubic, Yang et al, Xu et al and Nazzal et al algorithms



Original

Proposed

Bicubic



Yang et al

Xu et al

Nazzal et al

Figure 5.2: Visual Comparison of the Boat Image with the Bicubic, Yang et al, Xu et al and Nazzal et al algorithms

Chapter 6

CONCLUSION

6.1 Conclusion

In this thesis, single image super-resolution algorithm is proposed. Coupled K-SVD dictionary algorithm is used to design structured dictionaries in the wavelet domain which are compact and effectively directional, since they inherit the structural and directional properties of the respective wavelet subbands. The coupled K-SVD method improves the image reconstruction by enforcing the equality on the sparse representation coefficients of the LR and HR patches. Coupled K-SVD better exploits the persistence property of the wavelet coefficients and improves the coupling of the sparse representation coefficients of the HR and LR subbands. The proposed algorithm is experimentally tested both on the some Kodak set image and other benchmark images. The results indicate that proposed algorithm performs well in terms of PSNR and SSIM when compared with the state of art super resolution algorithms.

6.2 Future Task

In this thesis we have used the coupled K-SVD algorithm for SISR in wavelet domain. We have used the DWT for the analysis and synthesis of the signal. DWT takes into account the horizontal, vertical and diagonal details of the image, as an extension to this work we can also take more directional features into account for the dictionary learning and super-resolution such as anti-diagonal features.

REFERENCES

- [1] Elad, M., & Michal, A. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on* 15.12: 3736-3745.
- [2] Pati, Y.C., Ramin, R., & Krishnaprasad, P. S. (1993). Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. *Signals, Systems and Computer, Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE.
- [3] Daubechies, I. (2010). Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics* 63.1: 1-38.
- [4] Gorodnitsky, I.F., & Bhaskar D.R. (1997). Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *Signal Processing, IEEE Transactions on* 45.3: 600-616.
- [5] Efron, B. (2004). Least angle regression. *The Annals of statistics* 32.2: 407-499.
- [6] Tibshirani, R. (1994). Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*: 267-288, 1994.
- [7] Olshausen, B.A., & David J.F. (1996). Natural image statistics and efficient coding*. *Network: computation in neural systems* 7.2 : 333-339.

- [8] Engan, K., Sven O.A., & HakonHusoy, J. (1999). Method of optimal directions for frame design. *Acoustics, Speech, and Signal Processing, Proceedings*, IEEE.
- [9] Aharon, M., Elad, M. & Alfred B. (2006). K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *Signal Processing, IEEE Transactions on* 54.11: 4311-4322.
- [10] Xu, J., Chun, Q., & Zhiguo C. (2014). Coupled K-SVD dictionary training for super-resolution. *Image Processing (ICIP), IEEE International Conference*.
- [11] Yang, J., et al. (2012). Coupled dictionary training for image super-resolution. *Image Processing, IEEE Transactions on* 21.8 : 3467-3478.
- [12] Zepeda, J., & Fabrice L. (2006). Tandem Filter Bank-DFT Code for Bursty Erasure Correction. *Vehicular Technology Conference, VTC-2006 Fall. 64th*. IEEE.
- [13] Rath, G., & Christine G. (2004). Subspace-based error and erasure correction with DFT codes for wireless channels. *Signal Processing, IEEE Transactions on* 52.11: 3241-3252.
- [14] Elad, M., et al. (2005). Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Applied and Computational Harmonic Analysis* 19.3: 340-358.

- [15] Guleryuz, O.G. (2006). Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: theory. *Image Processing, IEEE Transactions on* 15.3: 539-554.
- [16] Elad, M., & Michal A. (2006). Image denoising via learned dictionaries and sparse representation." *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*.
- [17] Bryt, O., & Elad, M. (2008) Compression of facial images using the K-SVD algorithm. *Journal of Visual Communication and Image Representation* 19.4: 270-282.
- [18] Adams, M.D. (2001). The JPEG-2000 still image compression standard.
- [19] Yang, J., et al. (2008). Image super-resolution as sparse representation of raw image patches. *Computer Vision and Pattern Recognition. IEEE Conference*.
- [20] Zeyde, R., Elad, M. & Matan P. (2010). On single image scale-up using sparse-representations. *Curves and Surfaces*. Springer Berlin Heidelberg, 711-730.
- [21] Yang, J., et al. (2010). Image super-resolution via sparse representation. *Image Processing, IEEE Transactions on* 19.11: 2861-2873.
- [22] Lee, H., et al. (2006). Efficient sparse coding algorithms. *Advances in neural information processing systems*.

- [23] Candès, E.J. (2006). Compressive sampling." *Proceedings of the International Congress of Mathematicians: Madrid, August 22-30, invited lectures.*
- [24] Colson, B., Patrice M., & Gilles S. (2007). An overview of bilevel optimization. *Annals of operations research* 153.1: 235-256.
- [25] Yang, J., Kai Y., & Thomas H. (2010). Supervised translation-invariant sparse coding. *Computer Vision and Pattern Recognition (CVPR), IEEE Conference.*
- [26] Bruckstein, A.M., David L.D., & Elad, M. (2009). From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM review* 51.1: 34-81.
- [27] Dong, W., et al. (2011) Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *Image Processing, IEEE Transactions on* 20.7: 1838-1857.
- [28] Donoho, David L. "Compressed sensing." *Information Theory, IEEE Transactions on* 52.4: 1289-1306, 2006
- [29] Elad, M., & Irad Y. (2009). A plurality of sparse representations is better than the sparsest one alone. *Information Theory, IEEE Transactions on* 55.10: 4701-4714.
- [30] Farsiu, S., et al. (2004). Fast and robust multiframe super resolution. *Image processing, IEEE Transactions on* 13.10: 1327-1344

- [31] Hardie, R.C., Kenneth J.B., & Ernest E.A. (1997). Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *Image Processing, IEEE Transactions on* 6.12: 1621-1633.
- [32] Kodak lossless true color image suite, retrieved on June (2015): <http://r0k.us/graphics/kodak/>.
- [33] Levin, A., et al. (2012). Patch complexity, finite pixel correlations and optimal denoising." *Computer Vision–ECCV*. Springer Berlin Heidelberg, 73-86.
- [34] Mairal, J., et al. (2009) Supervised dictionary learning. *Advances in neural information processing systems*.
- [35] Mairal, J., Guillermo S., & Elad, M. (2012). Multiscale sparse image representation with learned dictionaries. *Image Processing, ICIP. IEEE International Conference*.
- [36] Nazzal, M., & Ozkaramanli, H. (2013) Improved single image super-resolution using sparsity and structured dictionary learning in wavelet domain. *Signal Processing and Communications Applications Conference (SIU)*, IEEE.
- [37] Rauhut, H., Karin S., & Pierre V. (2008). Compressed sensing and redundant dictionaries." *Information Theory, IEEE Transactions on* 54.5: 2210-2219.
- [38] Tipping, M.E., & Christopher M.B. (2006). Bayesian image super resolution. U.S. Patent No. 7,106,914. 12 Sep.

- [39] Daubechies, I. (1992). *Ten lectures on wavelets*. Vol. 61. Philadelphia: Society for industrial and applied mathematics.
- [40] Jia, K., Xiaogang W., & Xiaoou T. (2016). Image transformation based on learning dictionaries across image spaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.2: 367-380.
- [41] Nazzal, M., & Ozkaramanli, H. (2014). "Wavelet domain dictionary learning-based single image superresolution. *Signal, Image and Video Processing*.
- [42] Wang, S., et al. (2012). Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis." *Computer Vision and Pattern Recognition (CVPR), IEEE Conference*.
- [43] Zhang, K., et al. (2012). Multi-scale dictionary for single image super-resolution. *Computer Vision and Pattern Recognition (CVPR), IEEE Conference*.
- [44] Ophir, B., Michael L., & Elad, M. (2011). Multi-scale dictionary learning using wavelets. *Selected Topics in Signal Processing, IEEE Journal of* 5.5: 1014-1024.
- [45] Mallat, S., & Sifen Z. (1992). Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 7: 710-732.
- [46] Mallat, S., & Wen L.H. (1992). Singularity detection and processing with wavelets." *Information Theory, IEEE Transactions on* 38.2: 617-643.