

Semi-Coupled Dictionary Learning for Single Image Super Resolution

Zia Ullah

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

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
September 2016
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Mustafa Tümer
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. Erhan A. İnce
Supervisor

Examining Committee

1. Prof. Dr. Erhan A. İnce

2. Prof. Dr. Şener Uysal

3. Asst. Prof. Dr. Rasime Uyguroğlu

ABSTRACT

It has been demonstrated in the literature that patches from natural images could be sparsely represented by using an over-complete dictionary of atoms. In fact sparse coding and dictionary learning (DL) have been shown to be very effective in image reconstruction. Some recent sparse coding methods with applications to super-resolution include supervised dictionary learning (SDL), online dictionary learning (ODL) and coupled dictionary learning (CDL). CDL method assumes that the coefficients of the representation of the two spaces are equal. However this assumption is too strong to address the flexibility of image structures in different styles. In this thesis a semi-coupled dictionary learning (SCDL) method has been simulated for the task of single image super-resolution (SISR). SCDL method assumes that there exists a dictionary pair over which the representations of two styles have a stable mapping and makes use of a mapping function to couple the coefficients of the low resolution and high resolution data. While tackling the energy minimization problem, SCDL will divide the main function into 3 sub-problems, namely (i) sparse coding for training samples, (ii) dictionary updating and (iii) mapping updating. Once a mapping function T and two dictionaries D_H and D_L are initialized, the sparse coefficients for the two sets of data can be obtained and afterwards the dictionary pairs can be updated. Finally, one can update the mapping function T through coding coefficients and dictionary. During the synthesis process, a patch based sparse recovery approach is used with selective sparse coding. Each patch at hand is first tested to belong to a certain cluster using the approximate scale invariance feature, then the dictionary pairs along with the mapping functions of that cluster are used for its reconstruction. First the low resolution patches of sparse

coefficients are calculated by the selected dictionary which has low resolution, then, the patches of high resolution is obtained by using the dictionary which has high resolution and sparse coefficients of low resolution.

In this thesis, comparisons between the proposed and the CDL algorithms of Yang and Xu were carried out using two image sets, namely: Set-A and Set-B. Set A had 14 test images and Set-B was composed of 10 test images, however in Set-B 8 of the test images were selected from text images which are in grayscale or in colour tone. Results obtained for Set-A show that based on mean PSNR values Yang's method is the third best and Xu's method is the second best. The sharpness measure based SCDL method was seen to be 0.03dB better than Xu's method. For set-B only the best performing two methods were compared and it was seen that the proposed method had 0.1664dB edge over Xu's method. The thesis also tried justifying the results, by looking at PSD of individual images and by calculating sharpness based scale invariance percentage for patches that classify under a number of clusters. It was noted that when most of the frequency components were around the low frequency region the proposed method would outperform Xu's method in terms of PSNR.

For images with a wide range of frequency components (spread PSD) when the number of HR patches in clusters C2 and/or C3 was low and their corresponding SM-invariance ratios were also low then the proposed method will not be as successful as Xu's method.

Keywords: sparse representations, super-resolution, semi-coupled dictionary learning, power spectral density, scale invariance.

ÖZ

Literatürde gösterilmiştir ki doğal imgelerden alınan yamalar, boyutları resimdekinden daha büyük olan bir sözlükten alınacak öğeler ile seyrekçe betimlenebilmektedir. Aslında, seyrek betimleme ve sözlük öğrenme (DL) teknikleri imge geri çatımı için oldukça başarılıdır. Son zamanlarda önerilen ve yüksek çözünürlük uygulamaları bulunan seyrek betimleme metodları arasında güdümlü sözlük öğrenme (SDL), çevrim içi sözlük öğrenme (ODL) ve bağlantılı sözlük öğrenme (CDL) bulunmaktadır. Bu bağlantılı sözlük öğrenme yöntemi her iki alanın betimleyici katsayılarının eşit olduğunu varsaymaktadır. Ancak, bu varsayım farklı biçimleri bulunan imge yapılarını esnek bir şekilde anlatmak için çok güçlüdür. Bu tezde, sözü geçen SCDL tekniği SISR için uyarlanmış ve benzetim sonuçlarını yorumlamıştır. SCDL yöntemi her iki alanı birbirine bağlayan bir sözlük çifti bulunduğunu varsayan ve seyrek betimlemelerin birbiri ile eşleştirilebileceğini düşünüp, düşük ve yüksek çözünürlüklü verilerin katsayılarını birbirine eşleştirecek bir fonksiyon kullanan bir yöntemdir. Enerji enküçültme problemini çözmeye çalışan SCDL, problemi üç alt-probleme bölmektedir: (i) eğitim örnekleri ve seyrek betimleme, (ii) sözlük güncelleme ve (iii) eşleştirme fonksiyonu güncellemesi. Eşleştirme fonksiyonu W , D_H ve D_L sözlüklerinin iklendirmesini mütakip her iki veri seti için seyrek betimleme katsayıları elde edilmekte ve daha sonra sözlük çiftleri güncellenmektedir. Son olarak da, elde edilen katsayılar ve sözlüklerin kullanımı ile eşleştirme fonksiyonu güncellenmektedir. Sentezleme esnasında eldeki her yamanın hangi topağa ait olduğu belirlenmekte ve daha sonra eşleştirme fonksiyonu ve sözlükler kullanılarak yamanın geri çatımı gerçekleştirilmektedir.

Tezde, CDL tabanlı Yang ve Xu metodları ile önerilen yarı bağlantılı sözlük öğrenme yöntemi Set-A ve Set-B diye adlandırdığımız iki set imge kullanılarak kıyaslanmıştır. Set-A 14, Set-B ise 10 sınaama imgesinden oluşmaktadır. Set-B deki imgelerden sekizi gri tonlu ve renkli yazı sınaama imgeleri içermektedir.

Set-A kullanıldığında elde edilen ortalama PSNR değerleri Yang metodunun en iyi üçüncü sonucu ve Xu metodunun da en iyi ikinci sonucu verdiği gözlemlenmiştir. Netlik ölçüsü tabanlı SCDL yönteminden elde edilen ortalama PSNR değeri Xu ya göre 0.03dB daha yukardadır. Set-B kullanılırken sadece en iyi başarıyı göstermiş iki yöntem kıyaslanmıştır. Ortalama PSNR değerleri önerilen yöntemin Xu ya göre 0.166 dB daha iyi olduğunu göstermiştir. Tezde ayrıca güç spektral yoğunluğuna ve farklı topaklar altında düşen yamalar için netlike-ölçüsü değerleri aralıklarına bağlı olarak hesaplanan bir ölçü bağımsız yüzdelik hesabı kullanılarak elde edilen sonuçlar yorumlanmıştır. Görülmüştür ki çoğu frekans bileşenleri düşük frekanslı olduğu zaman önerilen yöntem ortalama PSNR baz alındığında Xu ve Yang'dan daha iyi sonuç sağlamaktadır.

Çok farklı frekans bileşenleri olan imgelerde ise (yayılmış PSD), C2 ve C3 topaklarına düşen HR yamaları ve bunlara denk gelen SM-değişimsizlik oranları düşük olduğu zaman önerilen yöntem en yakın rakibi Xu dan biraz daha kötü başarıyı göstermektedir.

Anahtar kelimeler: seyrek betimlemeler, yüksek çözünürlük, yarı-bağlaşımlı sözlük öğrenme, güç spectrum yoğunluğu, scale invariance.

DEDICATION

To my parents and my dear elder brother Mr. Misbah Ullah

ACKNOWLEDGEMENT

Many thanks to Almighty Allah, the beneficent and the gracious, who has bestowed me with strength, fortitude and the means to accomplish this work.

I extend my heartfelt gratitude to my supervisor Prof. Dr. Erhan A. İnce for the concern with which he directed me and for his considerate and compassionate attitude, productive and stimulating criticism and valuable suggestions for the completion of this thesis. His zealous interest, and constant encouragement and intellectual support accomplished my task.

I am also grateful to Prof. Dr. Hasan Demirel, the Chairman of the department of Electrical and Electronic engineering who has helped and supported me during my course of study at Eastern Mediterranean University.

Lastly I would like to acknowledge my friends Junaid Ahamd, Gul Sher, Muhammad Irfan Khan, Muhammad Sohail, Muhammad Yusuf Amin, Muhammad Waseem Khan, Sayed Mayed Hussain, Saif Ullah and Atta Ullah.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZ.....	v
DEDICATION.....	vii
ACKNOWLEDGEMENT.....	viii
LIST OF SYMBOLS & ABBREVIATIONS.....	xi
LIST OF TABLES.....	xiii
LIST OF FIGURE.....	xiv
1 INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Motivation.....	3
1.3 Contributions.....	4
1.4 Thesis Outline.....	4
2 LITERATURE REVIEW.....	6
2.1 Introduction.....	6
2.2 Frequency Domain Based SR Image Reconstruction.....	6
2.3 Regularization Based SR Image Reconstruction.....	9
2.4 Learning Based Super Resolution Techniques.....	10
2.5 Single Image Super Resolution Technique.....	11
2.6 Dictionary Learning Under Sparse Model.....	11
2.7 Single Image Super Resolution on Multiple Learned Dictionaries and Sharpness Measure.....	12
2.8 Application of Sparse Representation Using Dictionary Learning.....	13
2.8.1 Denoising.....	13
2.8.2 In-painting.....	13

2.8.3 Image Compression	14
3 SUPER-RESOLUTION	15
3.1 Introduction	15
3.2 Joint Sparse Coding.....	17
3.3 Coupled Dictionary Learning	18
3.4 K-SVD Based Dictionary Learning	20
4 THE PROPOSED SUPER RESOLUTION METHOD	23
4.1 Single Image Super Resolution	23
4.2.1 Proposed Dictionary Learning Approach.....	24
4.2.2 Training of Semi Coupled Dictionary Learning.....	26
4.2.3 Reconstruction of SR Image using Semi Coupled Dictionary Learning.....	29
5 SIMULATION RESULTS	32
5.1 Introduction	32
5.2 The Patch size and Number of Dictionary Atoms Effects on the Representation quality.....	33
5.3. Evaluation and Comparison of Proposed Algorithm	36
5.3 Quality of Super-Resolved Images.....	47
6 CONCLUSIONS AND FUTURE WORK	51
6.1 Conclusion.....	51
6.2 Future Work	53
REFERENCES.....	54

LIST OF SYMBOLS & ABBREVIATIONS

D	Dictionary
T	Sparsity
S	Down Sampling Operator
X	High Resolution Training Set
Y	Low Resolution Training Set
x	High Resolution Patch
y	Low Resolution Patch
μ_x	Average of X
σ_k	k^{th} Singular Value
$\sigma_{X,\hat{X}}$	Covariance of X, \hat{X}
$\ \cdot\ _0$	Zero Norm
$\ \cdot\ _F$	Frobenious Norm
$\ \cdot\ _2$	Euclidean Norm
HR	High Resolution
KSVD	K-Mean Cluster Value Decomposition
MP	Matching Pursuit
LARS	Least Angle Regression Stage-wise
LR	Low Resolution
SISR	Single Image Super Resolution
MSE	Mean Square Error
OMP	Orthogonal Matching Pursuit
SR	Super Resolution

PSD	Power Spectrum Density
SM	Sharpness Measure
MR	Mid Resolutions
PSNR	Peak to Signal Noise Ratio
SSIM	Structural Similarity Index
DFT	Discrete Fourier transform
SCDL	Semi-coupled Dictionary Learning
QCQP	Quadratically Constrained Quadratic Program
SI	Scale Invariance

LIST OF TABLES

Table 5.1: Average PSNR and SSIM results for (3×3) patch size and different number of dictionary atoms and samples.....	34
Table 5.2: Average PSNR and SSIM results for (5×5) patch size and different number of dictionary atoms and samples.....	35
Table 5.3: Average PSNR and SSIM results for (7×7) patch size and different number of dictionary atoms and samples.....	35
Table 5.4: Average PSNR and SSIM results for (9×9) patch size and different number of dictionary atoms and samples.....	36
Table 5.5: Proposed SCDL SISR method versus classic bi-cubic interpolator, Yang's algorithm and Xu's algorithm.	38
Table 5.6: Xu versus Proposed Method on test images in Set-B.....	42
Table 5.7 Scale Invariance of Regular Texture Images and Text images in Set-B ...	46
Table 5.8: Scale Invariance of Regular Texture Images and Text images in Set-A...47	

LIST OF FIGURE

Figure 2.1: Framework of Wavelet Based SR Image Reconstruction	8
Figure 2.2: The Interpolation based SR Image Reconstruction	9
Figure 2.3: Learning based SR Image Framework	10
Figure 4.1. Training Images Data Set	26
Figure 4.2: Atoms for Low Resolution Dictionaries.....	29
Figure 4.3: Atoms for High Resolution Dictionaries	31
Figure 5.1: The Power Spectral Density Plots of Different images in Set-A	41
Figure 5.2: Power Spectral Density Plots of Images in Set-B.....	45
Figure 5.3: SR images for test image 'AnnieYukiTim' from Set-A.....	48
Figure 5.4: SR images for test image 'Flower' from Set-A.....	48
Figure 5.5: SR images for test image 'Butterfly' from Set-A	49
Figure 5.6: SR images for test image 'Rocio' from Set-A	49
Figure 5.7: SR images for test image 'Starfish' from Set-A	50

Chapter 1

INTRODUCTION

1.1 Introduction

Super resolution (SR) has been one of the most active research topics in digital image processing and computer vision in the last decade. The aim of super resolution, as the name suggests, is to increase the resolution of a low resolution image. High resolution images are important for computer vision applications for attaining better performance in pattern recognition and analysis of images since they offer more detail about the source due to the higher resolution they have. Unfortunately, most of the times even though expensive camera system is available high resolution images are not possible due to the inherent limitations in the optics manufacturing technology. Also in most cameras, the resolution depends on the sensor density of a charge coupled device (CCD) which is not extremely high.

These problems can be overcome through the use of super-resolution techniques which are based on the idea that a combination of low resolution (noisy plus blurred) sequence of images of a scene can be used to generate a high resolution image or an image sequence. Most of the time, it is assumed that the low resolution images have resulted from resampling of a high resolution image. Accuracy of imaging models is quite vital for super-resolution processing. Incorrect modelling of motion can actually further degrade the image. The images which are super resolved give more information or give more minutiae about the original image and also have high pixel

density. A high resolution (HR) image can also be recovered from a single image via the use of a technique known as “Super resolution from single image”. Many researchers worked on this topic and gain some good quantitative and qualitative results as summarized by [1].

The authors of [1] have introduced the concept of patch sharpness as a measure and have tried to super-resolve the image using a selective sparse representation over a pair of coupled dictionaries. Sparsity has been used as regularize since the single-image super resolution process maintains the ill-posed inverse problem inside. The sparse representation is based on the assumption that a n -dimensional signal vector ($x \in R^n$) could be approximated as a linear combination of some selected atoms in a dictionary of k -atoms ($D \in R^{n \times k}$). Hence the approximation can be written as $x \approx D \cdot \alpha$ where α denotes a sparse coding vector mainly composed of zeros. The problem of determining this representation is generally referred to as the sparse coding process.

The single image super resolution algorithm is composed of two parts. The first part is the training stage where a set of dictionary pairs is learned and the second is the reconstruction stage in which the best dictionary pair is selected to sparsely reconstruct HR patches from the corresponding LR patches. Processing in the training stage would start by dividing each HR image into non-overlapping patches. Then these HR patches are reshaped and combined column-wise to form a HR training array. Next, a set of LR images are obtained by down sampling and blurring each HR image in the HR Training image set. Afterwards each LR image would be up-sampled to form mid-resolution (MR) images. Similar to the HR images the mid resolution images are divided into patches, vectorized and column-wise combined to

form a LR training array. Then for each patch in the LR training array the gradient profile magnitude operator is used to measure the sharpness of the corresponding patch in the MR image. Based on a set of selected sharpness measure (SM) intervals the patches would be classified into a number of clusters. The algorithm would then add the MR and the HR patches to the corresponding LR training set of the related cluster or to the corresponding HR training set. Finally, the training data in each cluster is used to learn a pair of coupled dictionaries (LR and HR dictionaries).

In the reconstruction stage from these learned dictionaries the SM value of each LR patch is used to identify which cluster it belongs to and the dictionary pair of the identified cluster is used for reconstructing the corresponding HR patch via the use of sparse coding coefficients. The high resolution patch is calculated by multiplying the high resolution clustered dictionary with the calculated coefficients.

In the literature researches have used coupled dictionary, semi coupled dictionary, coupled K-SVD dictionary to super resolve an image [3] [4]. In [5],[6] and [7] authors have made proposals to improve upon the results of Yang's approach. In this thesis we will super resolve the image with better semi-coupled dictionaries where the coupling improvement between LR and HR image coefficients are due to a new mapping function. The general approach is similar to what has been presented in [1][2][3][4],but the procedure is different and leads to improved results.

1.2 Motivation

As in [5], the objective of this thesis is to propose a new mapping function that would improve the coupling between the semi-coupled dictionaries in the single image super resolution technique and improve the peak-signal-to-noise ratio (PSNR)

values obtained for the super-resolved images. Clustering was used to regulate the intrinsic grouping of unlabelled data in a set. Since there is no absolute criteria for selecting the clusters the user can select his/her own criteria and carry out the simulations to see if the results would suit to their needs.

Various researchers have worked on coupled dictionary learning for single image super-resolution, but here we are doing coupling as well as mapping between HR and LR patches. For assessing the quality of the reconstructed HR image as in [9] the peak signal-to-noise-ratio (PSNR) and structural similarity index (SSIM) were used and the test image and the reconstructed image were compared.

1.3 Contributions

In this thesis we used the MATLAB platform to simulate semi-coupled dictionary learning and applied it to the problem of single image super resolution. Our contributions are twofold: firstly we have obtained results for the specific sharpness measure based mapping function that relates the HR and LR data and have compared them against results obtained from other state-of-the-art methods. Secondly we have tried to analyse the results. To this end we have used PSD and SM-invariance values to explain the results.

1.4 Thesis Outline

The remainder of this thesis is organized as follows: Chapter 2 provides a literature review on super-resolution techniques. This is followed in Chapter 3 by a detailed study of super resolution, single-image super resolution, coupled dictionary learning and K-SVD based DL methods. Chapter 4 gives details about our proposed SISR algorithm based on SCDL and its mapping function that relates two spaces to each other. Chapter 5 provides simulation results for the proposed method and compares

these results with those obtained from classical bi-cubic interpolator and two other state of the art methods. This chapter also tries to interpret the results. Finally, Chapter 6 provides conclusions and makes suggestions for future work.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

The aim of super resolution (SR) image reconstruction is to obtain a high resolution image from a set of images captured from the same scene or from a single image as stated in [28]. Techniques of SR can be classified into four main groups. The first group includes the frequency domain techniques such as [13, 14, 15], the second group includes the interpolation based approaches [16, 17], the third group is about regularization based approaches [19, 20], and the last group includes the learning based approaches [17, 18]. In the first three groups, a HR image is obtained from a set of LR images while the learning based approaches achieve the same objective by using the information delivered by an image database.

2.2 Frequency Domain Based SR Image Reconstruction

The first frequency-domain SR technique [13] was proposed by Huang and Tsai where they considered SR computation for noiseless LR images. Their idea was to transform the data of the low resolution images into the Discrete Fourier Transform (DFT) domain and then combine the data based on the relationship between the aliased DFT coefficients of the observed LR images and the unknown HR image. The combined data would finally be transformed back to the spatial domain. Even though frequency domain techniques have low computational complexity they are still insufficient for dealing with real-world applications. This is due to the fact that the frequency domain techniques require only a global displacement between

observed images and the linear blur function. Two examples to frequency-domain techniques include usage of the discrete cosine transform (DCT) to perform fast image deconvolution for SR image computation (proposed by Rhee and Kang) and the iterative expectation maximization (EM) algorithm presented by Wood et al. [21, 22]. The registration, blind deconvolution and interpolation operations are all simultaneously carried out in the EM algorithm.

In the literature many researchers have investigated the usage of wavelet transforms for addressing the SR problem. They have been motivated to use wavelets since the wavelet transform would provide multi-scale and resourceful information for the reconstruction of previously lost high frequency information [14]. As depicted in Fig. 2.1 wavelet transformation based techniques would treat the observed low-resolution images as the low-pass filtered sub-bands of the unknown wavelet-transformed high-resolution image. The main goal is to evaluate the finer scale coefficients. This is then followed by an inverse wavelet transformation to obtain the HR image. The low-resolution images are viewed as the representation of wavelet coefficients after several levels of decomposition. The final step include the construction of the HR image after estimating the wavelet coefficient of the $(N+1)^{\text{th}}$ scale and taking an inverse discrete transformation.

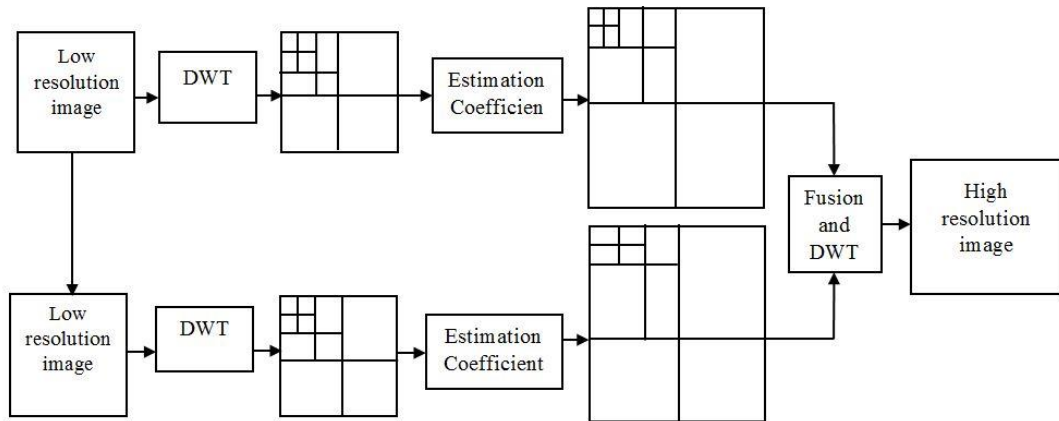


Figure 2.1: Framework of Wavelet Based SR Image Reconstruction [28]

The interpolation based SR image method involves projecting the low resolution images onto a reference image, and then fusing together the information available in the individual images. The single image interpolation algorithm cannot handle the SR problem well, since it cannot produce the high-frequency components that were lost during the acquisition process [14]. As shown by Fig 2.2, the interpolation based SR techniques are generally composed of three stages. These stages include:

- i. registration stage for lining up the LR input images,
- ii. interpolation stage for generating a higher resolution image, and the
- iii. de-blurring stage for enhancing the HR image obtained in step (ii).

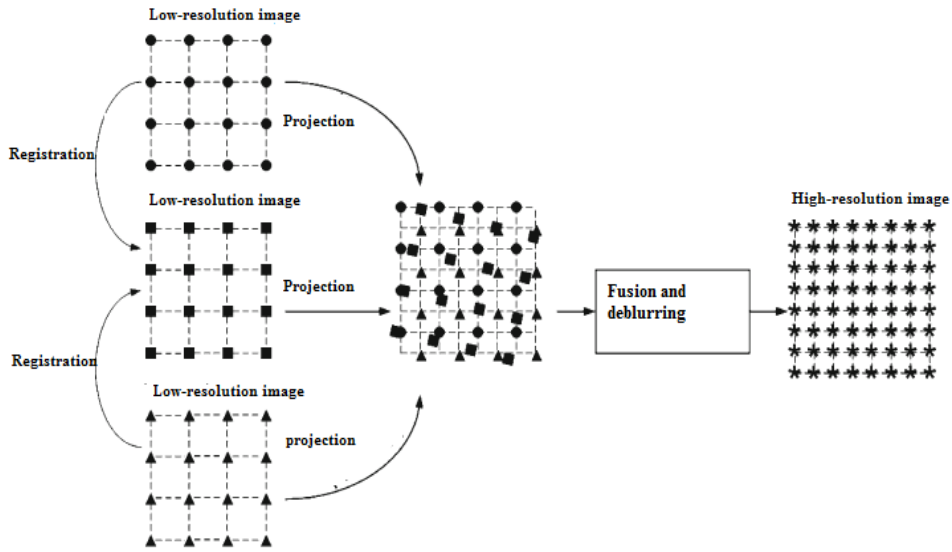


Figure 2.2: The Interpolation Based SR Image Reconstruction [28]

2.3 Regularization Based SR Image Reconstruction

In the literature we can find numerous regularization based SR image reconstruction methods [24, 25]. These methods have all been motivated by the fact that the SR computation was an ill-posed inverse problem. The aim of these regularization based SR methods is to incorporate the prior knowledge of the unknown high-resolution image into the SR process. According to the Bayesian point of view, the information that can be extracted from the low-resolution images about the unknown signal (HR image) is contained in the probability distribution of the unknown. Therefore the HR image can be estimated via some statistics of a probability distribution of the unknown high-resolution image, which can be established by applying Bayesian inference to exploit the information provided by both the low-resolution images and the prior knowledge on the unknown high-resolution image. The most popular Bayesian-based SR approaches include the maximum likelihood (ML) estimation [20] and the maximum a posterior (MAP) estimation [24] approaches.

2.4 Learning Based Super Resolution Techniques

Similar to the regularization based SR approaches the learning based SR techniques also try solving the ill-posed SR problem [26, 27]. The aim of these learning based methods is to enhance the high frequency content of the single LR input image by extracting the most likely high-frequency information from the given training image samples considering the local features of the input low-resolution image. For this end, Hertzman [26] had proposed a method in which the desired high-frequency information is recovered from a database of LR images. As can be seen from Fig. 2.3, Hertzman's method is made up of 2- stages. An off-line training stage and the SR image reconstruction stage. In the first stage image patches are used as ground truth for the generating LR patches through an image acquisition model proposed in [28]. The method would first collect pairs of LR patches and their corresponding HR patches. Later in the reconstruction stage, patches extracted from the input low-resolution images are compared with other patches stored in a database. Afterwards to get the HR image a similarity measurement criterion such as minimum distance is used to select the best matching patches.

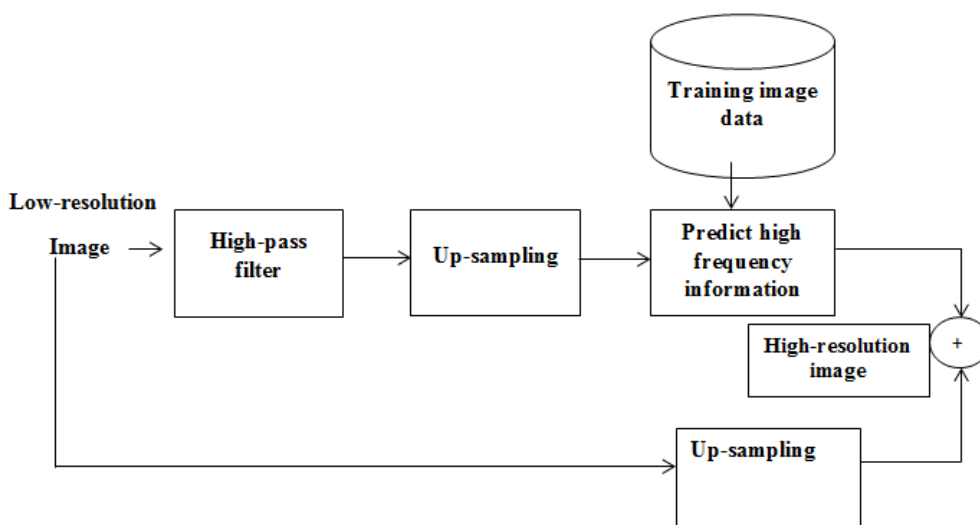


Figure 2.1: Learning based SR Image Framework

2.5 Single Image Super Resolution Technique

As stated in [3], the single image super resolution (SISR) technique only makes use of a single LR observation to construct a HR image. This method which had been proposed by Yang uses coupled dictionary training and involves patch wise sparse recovery. In Yang's approach to SR image reconstruction, the learned coupled dictionaries are used to relate the LR and HR image patch-spaces via sparse representation. The learning process makes sure that the sparse representation of a low-resolution image patch in terms of the low-resolution dictionary can be used to reconstruct a HR image patch with the dictionary in the high-resolution image patch space. The SISR method proposed by Yang, is a bi-level optimization problem where an L_1 -norm minimization is used among the constraints of the optimization.

2.6 Dictionary Learning Under Sparse Model

Dictionary learning for sparse representation is a very active research topic among researchers all around the world. As suggested in [3] by Yang et al., the main aim of sparse representation is to present the data in a meaningful way to capture the useful properties of signals with only a few coefficients that are nonzero (sparse). Dictionary learning using sparse representation had become a necessity due to the limited representation power of the orthogonal and bi-orthogonal DL methods. The sparse and redundant data modelling seeks the representation of signals as linear combinations of a small number of atoms from a pre-specific dictionary. Sparse coding is the name given to the method to discover such a good set of basis atoms. Given a training data $\{x_i\}_{i=1}^N$ the problem of dictionary learning is solving an optimization problem by minimizing the energy function that combines squared reconstruction errors and the L_1 sparsity penalties on the representations. Currently there is an abundance of DL methods and most of these methods use the L_0 and L_1

sparsity penalty measures. Some recent DL techniques include the method of optimal direction (MODs), the Recursive Least Squares Dictionary Learning (RLS-DLA) method, K-SVD dictionary learning and the online dictionary learning (ODL). MOD technique which was proposed by Engan et al. [30], employs the L_0 sparsity measure and K-SVD uses the L_1 penalty measure. These over-trained dictionaries have the advantage that they are adaptive to the signals of interest, which contributes to the state-of-the-art performance on signal recovery tasks such as in-painting [33] denoising [31] and super resolution [2],[9].

2.7 Single Image Super Resolution on Multiple Learned Dictionaries and Sharpness Measure

A new technique has recently been proposed in [1] by Yeganli et al. This new technique uses selective sparse representation over a set of coupled low and high resolution cluster dictionary pairs. Patch clustering and sparse model selection are based on a sharpness measure (SM) value obtained from the magnitude of the patch gradient operator. The patch sharpness measure used for HR and LR images is assumed to be independent to patch resolution. SM value intervals are used to cluster the LR input image patches. For each cluster, a pair of structured and compact LR and HR dictionaries is learned.

In the reconstruction stage, each LR patch is classified into a certain cluster using the SM value in hand. Afterwards, the corresponding HR patch is reconstructed by multiplying the cluster's HR dictionary with the sparse representation coefficients of the LR patch (as coded over the LR dictionary of the same cluster). Various numerical experiments using PSNR and SSIM have helped validate that the method of Faezeh is competitive with various state-of-the-art super-resolution algorithms.

2.8 Application of Sparse Representation Using Dictionary Learning

2.8.1 Denoising

In denoising of videos and images the sparse representation is also used [36]. In these problems the MAP approximation is expressed in which the priority of sparsity is on the base of given data. A solution is described for the MAP estimate where the sparse estimate in image block are overlapped and after that by taking average in every block the data is identified for denoising.

Let suppose X is noisy image of size $(R \times d)$ in which the overlapped patches x_k are being extracted and after that reshaped the patches to make a vector. Use a K-SVD algorithm for over-complete dictionary D , now use OMP for all patches to be sparsely coded. The finest atom d_k is considered as important part y_k such that the noisy part y_k is rejected.

$$N_i = D \cdot \alpha_i \quad (2.1)$$

At last reshaped all de-noise patches into 2-d patches. In overlapping patches take the average value pixel and obtained the de-noised image by merge them.

2.8.2 In-painting

In image processing the image in-painting application is very helpful now days. It is used for the filling of pixels in the image that are missed, in data transmission it is used to produce different channel codes and also it is useful for the removal of superposed text in manipulation or in road signs [34].

Now let suppose consider an image patch $x = [x_a^t \ x_b^t]^t$ that is made from the two sub-vectors where the sub-vector x_a is known for feasible pixel and x_b have a missing

pixels in image in-painting so there is a method proposed for the estimation of the missing sub-vector x_b by Guleryuz [8] in which missing pixels occupy combination in orthonormal bases for compression. For the compressibility of x signal means there is a sparse vector α exists that gives D such as, $x = D\alpha$. now let suppose Δ_a, Δ_m and Δ_n are the diagonal matrices with diaognalized value of 1/10 to describe the α non zero entries, so the i^{th} estimation for the x is given as;

$$\hat{x}^i = \begin{pmatrix} x_a \\ \hat{x}_b^i \end{pmatrix} = \Delta_a \hat{x}^{i-1} + \Delta_m (D \Delta_n D^t) \hat{x}^{i-1} \quad (2.2)$$

2.8.3 Image Compression

For image compression the over-complete dictionaries are used for sparsely representing the input signal to get good compression. An example of such approach is the work which is introduced by Bryt and Elad based on the learned K-SVD dictionary [20]. In their approach, they use some pre-specified face templates. Face templates which are not overlapped with the others, are used in order to specify a class of the signals. And then they are represented employing the corresponding K-SVD dictionary.

Chapter 3

SUPER-RESOLUTION

3.1 Introduction

In this chapter an introduction to super-resolution techniques is presented. Image super resolution (SR) techniques aim at estimating a high-resolution (HR) image from one or several low resolution (LR) observation images. SR carried out using computers mostly aim at compensating the resolution loss due to the limitations of imaging sensors. For example cell phone or surveillance cameras are the type of cameras with such limitations.

Methods for SR can be broadly classified into two families: (i) The traditional multi-image SR, and (ii) example based SR. Traditional SR techniques mainly need multiple LR inputs of the same scene with sub-pixel separations (due to motion). A serious problem with this inverse problem is that most of the time there are insufficient number of observations and the registration parameters are unknown. Hence, various regularization techniques have been developed to stabilize the inversion of this ill posed problem. If enough LR images are available (at subpixel shifts) then the set of equations will become determined and can be solved to recover the HR image. Experiments have showed that traditional methods for SR would lead to less than 2 percent increase in resolution [37].

In example based SR, the relationship between low and high resolution image patches are learned from a database of low and high resolution image pairs and then applied to a new low-resolution image to recover its most likely high-resolution version. Higher SR factors have often been obtained by repeated applications of this process, however example based SR does not guarantee to provide the true (unknown) HR details. Although SR problem is ill-posed, making precise recovery impossible, the image patch sparse representation demonstrates both effectiveness and robustness in regularizing the inverse problem.

Suppose we have a LR image Y that is obtained by down sampling and blurring a HR image X . Given an over-complete dictionary of K atoms ($K > n$), signal $X \in \mathbb{R}^n$ can be represented as a sparse linear combination with respect to D . In other words, the signal X can be written as $X = D \alpha_H$, where $\alpha_H \in \mathbb{R}^K$ is a vector with very few nonzero entries. In practice, one may have access to only a small set of the observations from X , say Y :

$$Y = LX = LD\alpha_H \quad (3.1)$$

Here L represents the combined effect of down sampling and blurring. $L \in \mathbb{R}^{k \times n}$ with $k < n$ and is known as a projection matrix.

Assuming that in the context of super-resolution x denotes a HR image patch, while y is its LR counterpart it is possible to sparsely represent the HR and LR patches using the over-complete dictionary D_H and D_L :

$$x \approx D_H \alpha_H, \quad y \approx D_L \alpha_L$$

By supposing that $D_L \approx LD_H$, then

$$\mathbf{y} \approx \mathbf{L}\mathbf{D}_H\alpha_H \approx \mathbf{D}_L\alpha_H \quad (3.2)$$

(3.2) describes the relationship between the sparse representation coefficient of LR and HR patches and says that they are approximately same, i.e. $\alpha_H \approx \alpha_H$. Therefore, given a LR patch \mathbf{y} , one can reconstruct the corresponding HR patch as:

$$\hat{\mathbf{x}} \approx \mathbf{D}_H\alpha_L. \quad (3.3)$$

Vector selection for sparse representation of \mathbf{y} can be formulated as:

$$\min_{\alpha_0} \|\mathbf{y} - \mathbf{D}_L\alpha_0\|_2, \text{ s.t. } \|\alpha_0\|_0 < T \quad (3.4)$$

Where, T is a threshold which is used to control the sparseness. $\|\cdot\|_2$ and $\|\cdot\|_0$ respectively represent the L2 and L0 norms.

To represent a signal, a well-trained dictionary and a sparse linear combination of the dictionary atoms is required. First, the initial dictionary is used to sparsely represent the observation and afterwards the dictionary is updated using the sparse representation for the given data.

3.2 Joint Sparse Coding

Unlike the standard sparse coding, the joint sparse coding involves learning two dictionaries \mathbf{D}_x and \mathbf{D}_y for two coupled feature spaces, χ and Υ . The two spaces are tied by a mapping function F . The sparse representation of $x_i \in \chi$ in terms of \mathbf{D}_x should be the same as that of $y_i \in \Upsilon$ in terms of \mathbf{D}_y , where $y_i = F(x_i)$. Yang et al. [3] addressed this problem by generalizing the basic sparse

$$\mathbf{z}_i = \operatorname{argmin}_{\alpha_i} \|\mathbf{y}_i - \mathbf{D}_y \alpha_i\|_2^2 + \mu \|\alpha_i\|_1, \forall i=1, 2, \dots, N \quad (3.7)$$

$$\mathbf{z}_i = \operatorname{argmin}_{\alpha_i} \|\mathbf{x}_i - \mathbf{D}_x \alpha_i\|_2^2 \quad \forall i=1, 2, \dots, N$$

Here $\{\mathbf{x}_i\}_{i=1}^N$ are the latent space samples, $\{\mathbf{y}_i\}_{i=1}^N$ are the samples from observation space and $\{\mathbf{z}_i\}_{i=1}^N$ denotes the sparse representations. We recall that $\mathbf{y}_i = F(\mathbf{x}_i)$ where $F(\cdot)$ is the mapping function.

Signal recovery from coupled spaces is similar to compressed sensing [32]. In compressed sensing there is a linear random projection function F . Dictionary \mathbf{D}_x is chosen to be a mathematically defined basis and \mathbf{D}_y is obtained directly from \mathbf{D}_x with the linear mapping. For more general scenarios where the mapping function F is unknown and may be non-linear the compressive sensing theory cannot be applied. For an input signal y , the recovery of its latent signal x consists of two consecutive steps: (i) find the sparse representation \mathbf{z} of y in terms of \mathbf{D}_y according to (3.7), and then estimate the latent signal, $\mathbf{x} = \mathbf{D}_x \mathbf{z}$.

To minimize the recovery error of \mathbf{x} , we define the following loss term:

$$L(\mathbf{D}_x, \mathbf{D}_y, \mathbf{x}, \mathbf{y}) = \frac{1}{2} \|\mathbf{D}_x \mathbf{z} - \mathbf{x}\|_2^2 \quad (3.8)$$

The optimal dictionary pair $\{\mathbf{D}_x^*, \mathbf{D}_y^*\}$ can be obtained by minimizing (3.8) over the training signal pairs as:

$$\begin{aligned} & \min_{\mathbf{D}_x, \mathbf{D}_y} \frac{1}{N} \sum_{i=1}^N L(\mathbf{D}_x, \mathbf{D}_y, \mathbf{x}_i, \mathbf{y}_i) \\ & \text{s.t. } \mathbf{z}_i = \operatorname{argmin}_{\alpha} \|\mathbf{y}_i - \mathbf{D}_y \alpha\|_2^2 + \mu \|\alpha\|_1, \forall i = 1, 2, \dots, N \\ & \|\mathbf{D}_x(:, k)\|_2 \leq 1, \|\mathbf{D}_y(:, k)\|_2 \leq 1, \forall k = 1, 2, \dots, N \end{aligned} \quad (3.9)$$

Objective function in (3.09) is nonlinear and highly nonconvex. To minimize it one must alternatively optimize over \mathbf{D}_x and \mathbf{D}_y while keeping the other fixed. When \mathbf{D}_y is fixed, sparse representation \mathbf{z}_i can be determined for each \mathbf{y}_i with \mathbf{D}_y , and the problem reduces to:

$$\begin{aligned} \min_{\mathbf{D}_x} \sum_{i=1}^N \frac{1}{2} \|\mathbf{D}_x \mathbf{z}_i - \mathbf{x}_i\|_2^2 \\ \text{s.t. } \mathbf{z}_i = \underset{\alpha}{\text{argmin}} \|\mathbf{y}_i - \mathbf{D}_y \alpha\|_2^2 + \mu \|\alpha\|_1, \forall i = 1, 2, \dots, N \\ \|\mathbf{D}_x(:, k)\|_2 \leq 1, \forall k = 1, 2, \dots, k \end{aligned} \quad (3.10)$$

This is a quadratically constrained quadratic programming problem that can be solved using conjugate gradient descent algorithm [35]. Minimizing the loss function of (3.10) over \mathbf{D}_y is a highly nonconvex bi-level programming problem as stated in [37] and this problem can be solved through the use of the gradient descent algorithm [38].

3.4 K-SVD Based Dictionary Learning

In vector quantization (VQ), a codebook \mathbf{C} that includes K codewords is used to represent a wide family of signals $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^N$ by a nearest neighbour assignment. The VQ problem can be mathematically described as:

$$\min_{\mathbf{C}, \mathbf{X}} \{\|\mathbf{Y} - \mathbf{C}\alpha\|_F^2\} \quad \text{s.t. } \forall i \mathbf{x}_i = \mathbf{e}_k \quad (3.11)$$

The K-Means algorithm [38] is an iterative method, used for designing the optimal codebook for VQ. In each iteration of K-Means there are two steps. The first step is the sparse coding that essentially evaluates α by mapping each signal to its closest atom in \mathbf{C} , and the second step is the updating of the codebook, changing sequentially each column c_i in order to better represent the signals mapped to it.

Sparse representation problem can be viewed as a generalization of (3.11), in which we allow each input signal to be represented by a linear combination of codewords, which we now call dictionary elements or atoms. Hence, (3.11) can be re-written as:

$$\min_{\mathbf{D}, \boldsymbol{\alpha}} \{ \|\mathbf{Y} - \mathbf{D}\boldsymbol{\alpha}\|_F^2 \} \quad s. t. \quad \forall i \quad \|x_i\|_0 \leq T \quad (3.12)$$

In K-SVD algorithm we solve (3.12) iteratively in two steps parallel to those in K-Means. In the sparse coding step, we compute the coefficients matrix $\boldsymbol{\alpha}$, using any pursuit method, and allowing each coefficient vector to have no more than T non-zero elements. In second step dictionary atoms are updated to better fit the input data. Unlike the K-Means generalizations that freeze $\boldsymbol{\alpha}$ while determining a better \mathbf{D} the **K-SVD** algorithm changes the columns of D sequentially and also allows changing the coefficients as well. Updating each d_k has a straight forward solution, as it reduces to finding a rank-one approximation to the matrix of residuals;

$$E_k = \sum_{j \neq k} d_j \alpha^j$$

Where α^j denote the j-th row in the coefficient matrix $\boldsymbol{\alpha}$. E_k can be restricted by choosing only the columns corresponding to those elements that initially used d_k in their representation. This will give E_k^R . Once E_k^R is available SVD decomposition is applied as:

$$E_k^R = U\Delta V^T$$

Finally we will update d_k and α_k^R as:

$$d_k = \mathbf{u}_1, \quad \alpha_k^R = \Delta(1,1) \cdot \mathbf{v}_1$$

While K-Means applies K mean calculations to evaluate the codebook, the K-SVD obtains the updated dictionary by K-SVD operations, each producing one column.

For coupled dictionary learning (3.12) will first be replaced by:

$$\begin{aligned}
 & \min_{D_x, D_y, \{\alpha_i\}} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{D}_x \mathbf{a}_i\|_2^2 + \|\mathbf{y}_i - \mathbf{D}_y \mathbf{a}_i\|_2^2 \\
 & \text{s.t } \|\mathbf{a}_i\|_0 \leq T, \|\mathbf{d}_x^r\|_2 \leq 1, \|\mathbf{d}_y^r\|_2 \leq 1 \\
 & \quad i = 1, 2, \dots, N, \quad r = 1, 2, \dots, n
 \end{aligned} \tag{3.13}$$

Where, \mathbf{d}_x^r and \mathbf{d}_y^r are the r -th atom of \mathbf{D}_x and \mathbf{D}_y respectively.

Afterwards, the two step K-SVD will be implemented.

Chapter 4

THE PROPOSED SUPER RESOLUTION METHOD

4.1 Single Image Super Resolution

This idea of Single Image Super Resolution (SIRS) by sparse representation rides on a very special property of sparse representation. According to model of the sparse land, a vector X_H can be denoted through a dictionary D_H and sparse representation vector α_H . Let X_H be the high resolution image patch and D_H and α_H by the corresponding dictionary and sparse coefficient vector. In the same way let X_L be the low resolution image patch, D_L and α_L be the corresponding LR dictionary and sparse representations vector, then X_H and X_L can be represented as:

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

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

Here, the X_L patches are generated by applying down sampling and blurring on HR images and then extracting the LR patches from these images. Now consider the invariance property of this sparse representation due to the resolution blur. Given the trained LR and HR dictionaries one can estimate HR patches from LR patches using the HR dictionary and calculate the LR coefficients.

$$\alpha_H \approx \alpha_L \quad (4.3)$$

$$X_H \cong D_H \quad (4.4)$$

This is a very basic idea of a sophisticated SR process. In [2] [5], Yang has proposed a mechanism of coupled dictionary learning in which the authors learn the dictionaries in the coupled space instead of using a single dictionary for each space. Here, the HR and LR data are concatenated to form a joint dictionary learning problem. Further, for the sparse coding stage authors suggest a joint by alternate mechanism and use it for both dictionaries in the dictionary update stage.

In [1], Yeganli and Nazzal have proposed a sharpness based clustering mechanism to divide the joint feature space into different classes. The idea here is that the variability of signal over a particular class is less as compared to its variability in general. They designed pairs of clustered dictionaries using the coupled learning mechanism of Yang et al. [2], and attained state of the art results.

Motivated from [1], in this thesis a new method is proposed for semi-coupled dictionary learning through feature space learning. Moreover we introduce a different sharpness feature than the one introduced in [1] for classifying patches into different clusters.

4.2.1 Proposed Dictionary Learning Approach

Before discussing the details of the proposed dictionary learning approach, we provide here first the clustering criteria and data preparation for the semi-coupled dictionary learning. The sharpness criteria used for the classification of patches into different clusters was as follows:

$$S(x, y) = \|I(x, y) - \mu_A(x, y)\|_1 \quad (4.5)$$

Equation (4.5) says that the sharpness at any pixel location (x, y) is equal to the L_1 distance between $I(x, y)$ and the mean $\mu_A(x, y)$ of its 8-adjacent neighbours. In this way we calculate a sharpness value for a particular patch. And assuming the invariance of this measure due to the resolution blur we consider only those patches which satisfy this invariance. Based on these criteria three clusters were created and three different class dependent HR and LR dictionaries are learned. We extract the patches from the same spatial locations for both HR and LR data.

For the dictionary learning process we have used a subset of the training image data set provided by Kodak [7]. In our subset we use 69 training images randomly selected from the Kodak set. Figure 4.1 depicts some of the images that are in this subset.



Figure 4.1. Training Images Data Set

4.2.2 Training of Semi Coupled Dictionary Learning

Let us consider the HR training image set as \mathbf{X} . We first down sample and blur the HR images by bi-cubic interpolation and generate the low resolution training image set as \mathbf{Y} . Now we sample the patches after the LR and HR training set image from each image from same spatial locations. Regarding the clustering of data while sampling the patches from LR and HR training data we test the sharpness value of HR and LR and if this value is same we cluster them into the same cluster. Then HR and LR 2D patches were converted into column vector and stacked column wise into

a cluster matrix. One more thing to note here is that for the LR training patches, we first calculate the gradient features of the patches as done by most SR algorithms before concatenating them into cluster matrix. Let U^y be the high resolution HR patches matrix where y represent the cluster number. In the same way let V^y be the low resolution patches matrix. The joint coupled dictionary learning process can be expressed as;

$$\begin{aligned} \min \{ D_u^y, D_v^y, f(\cdot) \} & \|U^y - D_u^y \Lambda_u^y\|_F^2 + \|V^y - D_v^y \Lambda_v^y\|_F^2 + \\ & \beta \|\Lambda_v^y - T^y \Lambda_u^y\|_F^2 + \varphi_\mu \|\Lambda_u^y\|_1 + \varphi_\nu \|\Lambda_v^y\|_1 + \varphi_m \|T\|_F^2 \\ \text{s.t. } & \|D_{u,i}^y\|_{l_2} \leq 1, \|D_{v,i}^y\|_{l_2} \leq 1. \end{aligned} \quad (4.6)$$

Here $\beta, \varphi_\mu, \varphi_\nu, \varphi_m$ are the regularization terms and d_{ui}^y and d_{vi}^y are the respective cluster dictionary. The problem posed in the above equation is solved in three ways, first it is solved for the sparse representation coefficients and keeping constant the dictionaries and mapping function. Then it is solved for the dictionaries while keeping the sparse representation and mapping functions constant. Finally it is solved for mapping function and keeping constant the sparse representation and dictionaries.

First given the HR and LR cluster training data for each cluster we initialize a dictionary and mapping matrix. Given all these dictionaries to the sparse representation problem and can be formulated as.

$$\min \{ \Lambda_u^y \} \|U^y - D_u^y \Lambda_u^y\|_F^2 + \beta \|\Lambda_v^y - T_u \Lambda_u^y\|_F^2 + \varphi_u \|\Lambda_u^y\|_1 \quad (4.7)$$

$$\min \{ \Lambda_v^y \} \|V^y - D_v^y \Lambda_v^y\|_F^2 + \beta \|\Lambda_u - T_v \Lambda_v^y\|_F^2 + \varphi_v \|\Lambda_v^y\|_1 \quad (4.8)$$

The problem posed above is a typical Lasso problem or one can say a vector selection problem. There are many algorithms in the literature that can solve this

problem such as LARS [11]. After finding the sparse coefficients of the HR and LR training data the dictionaries must be updated. This problem can be formulated as:

$$\begin{aligned} \text{Min } \{D_u^y, D_v^y\} & \|U^y - D_u^y \Lambda_v^y\|_F^2 + \|V^y - D_v^y \Lambda_u^y\|_F^2 & (4.9) \\ \text{s.t. } & \|D_{u,i}^y\|_{l_2} \leq 1, \quad \|D_{v,i}^y\|_{l_2} \leq 1 \end{aligned}$$

The problem posed in (4.9) is a quadratically constrained program (QCQP) problem and can be solved as described in [3]. The third step in the dictionary learning process is the updating of the mapping matrix. This problem can be stated as:

$$\text{Min } \{T^y\} \|\Lambda_u^y - T^y \Lambda_v^y\|_F^2 + (\varphi_m/\beta) \|T^y\|_F^2 \quad (4.10)$$

(4.10) shows a ridge regression problem and has been analytically solved. The mapping function T^y would be equal to:

$$T^y = \Lambda_u^y \Lambda_v^{yT} (\Lambda_u^y \Lambda_v^{yT} + (\varphi_m/\beta)I)^{-1} \quad (4.11)$$

where, I denotes the identity matrix.

This proposed dictionary algorithm is summarized below.

Algorithm 1 The Training Phase

Input: HR Training Image Set.

- Perform down-sampling and blurring on HR images to get the LR images.
 - Do bi-cubic interpolation on LR images to get the transformed MR images.
 - Extract patches from HR and MR images and classify them using the sharpness value into clusters to get the HR and LR training data matrices.
 - For each directional cluster having HR and LR training data matrices **X and**
-

Y do:

- Initialize dictionary pairs \mathbf{D}_H and \mathbf{D}_L for each cluster.
- Initialize mapping functions \mathbf{T}_H and \mathbf{T}_L in every cluster.
- Fix the further variables; update Λ_H and Λ_L through Eq (3.6).
- Fix other variables; update \mathbf{D}_H and \mathbf{D}_L by Eq.(3.6) .
- Fix other variables; update \mathbf{T}_H and \mathbf{T}_L by Eq.(3.6) .

Output: \mathbf{D}_H , \mathbf{D}_L , \mathbf{T}_H and \mathbf{T}_L for each cluster.

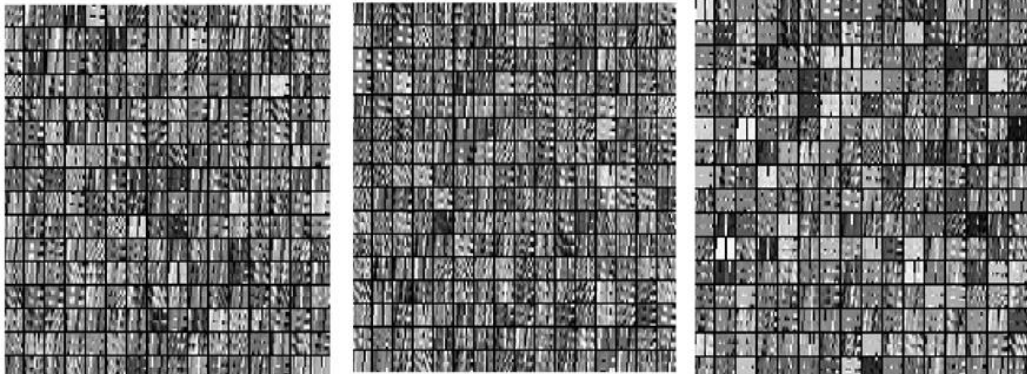


Figure 4.2: Atoms for Low Resolution Dictionaries

Figure 4.2 shows sample atoms of the dictionaries learned for the three clusters. On the left are the atoms that are not so sharp and on the right the ones that are sharpest. These are the atoms that will be used for the HR patch estimation during the reconstruction phase.

4.2.3 Reconstruction of SR Image using Semi Coupled Dictionary

Learning

During the image reconstruction stage an LR image given we first convert 2D image into column matrix for HR image reconstruction. One thing to remember here is that we apply full overlap and each LR patch is first tested by its sharpness value to

decide the dictionary pair for its reconstruction. In the patch wise sparse recovery process we first calculate the sharpness value of the LR patch at hand. From this sharpness value we find the cluster to which this patch belongs to. Given the cluster dictionaries and mapping matrix we calculate the sparse coefficient of this LR patch using the LR dictionaries and the mapping matrix by the same equation used during the training stage. Now after finding the sparse coefficients matrix we first multiply the sparse coefficients by the training matrix and then use the dictionary of HR along by multiplied sparse coefficients to estimate an HR patch. After estimating all HR patches. We go from the vector domain to the 2-D image space by using the merge method of Yang et al [2]. At the end we have our HR estimate image.

Algorithm 2 The Reconstruction Phase

Input: LR Test Image and Dictionary Pairs with Mapping Functions.

- Up-convert the LR image by bi-cubic interpolation.
- Extract patches from this transformed LR image using full overlap.
- Use the mapping function and dictionary pair of each cluster to recover the HR patch.
- For each LR patch test its sharpness value and decide the cluster it belongs to.
- Use the selected cluster LR dictionary and mapping to get the sparse coefficients of LR patch.
- Use the selected cluster HR dictionary and mapping to recover the HR patch using the sparse invariance property.

-
- Reshape and merge the recovered HR patches to get the approximate HR image.

Output: A HR image estimate.

Figure 4.3 depicts the atoms of the dictionaries for the reconstructed HR patches of the three clusters.

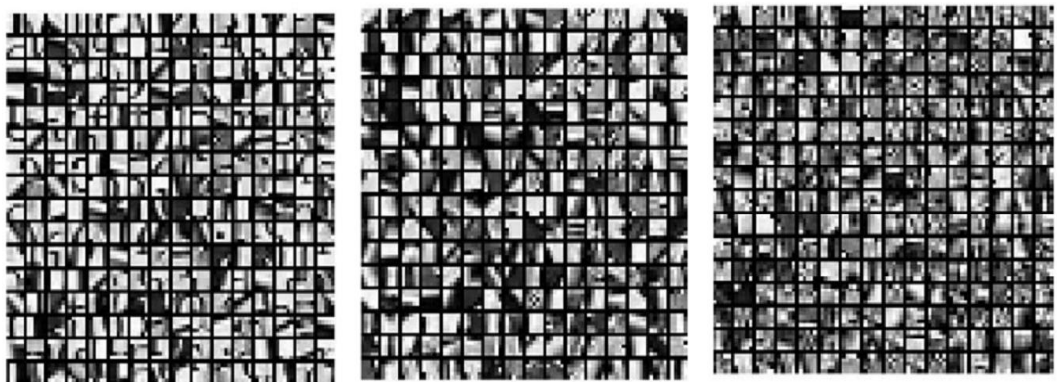


Figure 4.3: Atoms for High Resolution Dictionaries

Chapter 5

SIMULATION RESULTS

5.1 Introduction

In this chapter the performance of the proposed single image SR (SISR) algorithm based on semi-coupled dictionary learning (SCDL) and a sharpness measure (SM) has been evaluated and the results compared against the performance of the classic bi-cubic interpolation and two other state-of-the-art algorithms namely: techniques proposed by Yang [3] and Xu et al [4]. The dictionary of the proposed SISR algorithm was trained using 69 images from the Kodak image database [12], and a second benchmark data set [13]. After training the dictionary of the proposed SISR method, performance estimation of the reconstructed images were carried out using two sets of test images, namely: Set-A and Set-B. Metrics used in comparisons included the PSNR and SSIM.

Peak signal-to-noise ratio (PSNR) represents the ratio between the maximum possible value (power) of signal and power of distorting noise that disturbs the quality of its representation. Because various signals have very wide dynamic range PSNR is usually expressed in terms of the logarithmic decibel scale and is computed as follows:

$$PSNR(X, \hat{X}) = 10 \log_{10} \frac{255}{MSE(X, \hat{X})}, \quad (5.1)$$

Where, X is the original high resolution (HR) image and \hat{X} is the reconstructed image. Mean square error it's the error between (X, \hat{X}) which is defined as:

$$MSE(X, \hat{X}) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - \hat{x}_{ij})^2 \quad (5.2)$$

The luminance surface of an object is being observed as the product of illumination and the reflectance, but the structures of the objects in the scene are independent of the illumination. Consequently, to explore the structural information in the reconstructed images, we used a second metric which known as Structural Similarity index (SSIM). SSIM which compares local patterns of pixel intensities that have been normalized for luminance and contrast is defined as:

$$SSIM(X, \hat{X}) = \frac{(2\mu_X \mu_{\hat{X}} + C_1) (2\sigma_{X\hat{X}} + C_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + C_1) (\sigma_X^2 + \sigma_{\hat{X}}^2 + C_2)} \quad (5.3)$$

In (5.3), $\mu_X, \mu_{\hat{X}}$ are the average of the original image X and noisy image \hat{X} and $\sigma_X^2, \sigma_{\hat{X}}^2$ are noisy and original image variance where covariance for X, \hat{X} is $\sigma_{X\hat{X}}$. This covariance is computed as:

$$\sigma_{x, \hat{x}} = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - \mu_x)(\hat{x}_i - \mu_{\hat{x}}) \quad (5.4)$$

5.2 The Patch size and Number of Dictionary Atoms Effect on the Representation quality

In sparse representation, the dictionary redundancy is an important term as stated in [3]. In patch-based sparse representation dictionary redundancy is loosely defined as

a ratio for the dictionary atoms to patch size. Larger patch sizes mean larger dictionary atoms and when larger patches are used this will help better represent the structure of the images. However, when patches of image are large the learning process needs significantly more training images and inherently leads to an increase in the computational complexity of the dictionary learning (DL) and sparse coding processes. To see the effect of patch sizes on representation quality we used patch sizes of (3×3) , (5×5) , (7×7) , and (9×9) . Firstly, we checked the four patch sizes with 256 dictionary atoms. 10,000 sample patches were extracted from a set of natural images. It was made sure that the test images were not from the training images set. To assess the effect of increased dictionary atoms on performance, we have also used the four patch sizes together with 600 dictionary atoms taking 40,000 samples and 1200 dictionary atoms taking 40,000 samples. From the results obtained we noted that increasing the size of patch (hence the number of atoms) will lead to higher PSNR but the computations time will also increase.

Table 5.1: Average PSNR and SSIM results for (3×3) patch size and different number of dictionary atoms and samples.

3×3	Average PSNR	SSIM
Dictionary atoms: 256 Samples: 1000	27.77348 dB	0.853831
Dictionary atoms: 600 Samples: 40,000	27.84722	0.862038
Dictionary atoms: 1200 Samples: 40,000	27.89209	0.862745

Table 5.2: Average PSNR and SSIM results for (5×5) patch size and different number of dictionary atoms and samples.

5×5	Average PSNR	SSIM
Dictionary atoms: 256 Samples: 10,000	27.92892	0.862715
Dictionary atoms: 600 Samples: 40,000	28.02727	0.864475
Dictionary atoms: 1200 Samples: 40,000	28.07202	0.865774

Table 5.3: Average PSNR and SSIM results for (7×7) patch size and different number of dictionary atoms and samples.

7×7	Average PSNR	SSIM
Dictionary atoms: 256 Samples: 10,000	27.96074	0.86324
Dictionary atoms: 600 Samples: 40,000	28.04904	0.864984
Dictionary atoms: 1200 Samples: 40,000	28.09911	0.865902

Table 5.4: Average PSNR and SSIM results for (9×9) patch size and different number of dictionary atoms and samples.

9×9	Average PSNR	SSIM
Dictionary atoms: 256 Samples: 10,000	27.93783	0.862745
Dictionary atoms: 600 Samples: 40,000	28.0251	0.864353

5.3. Evaluation and Comparison of Proposed Algorithm

For evaluations, the proposed SISR algorithm is compared with two current leading super-resolution algorithms and the ever green bi-cubic interpolation technique. Throughout the simulations the same set of general simulation parameters were assumed for each algorithm and in addition for the bi-cubic technique we have used Matlab’s “**imresize**” function for upsampling.

The comparisons are made with Yang et al. [3] which is considered as a baseline algorithm and with Xu et al. [4] which utilizes a similar kind of dictionary learning and super-resolution approach as in [3], but the dictionary update is done using the K-SVD technique. Comparisons were carried out using two image sets, namely: Set-A and Set-B. Set A had 14 test images from the Kodak set [12] and Set-B was composed of 10 test images, six from the Flickr image set and 4 from internet sources. However in Set-B 8 of the test images were selected from text images.

Individual and mean PSNR and SSIM values for test images in Set-A is depicted in Table 5.5. This is clear from table that for Set-A, the lowest mean PSNR

performance belongs to the classical bi-cubic interpolator. Yang's method is the third best and Xu's method is the second best which is 0.03dB behind the average PSNR value for the proposed method. After a careful look at the individual PSNR values we noticed that the proposed semi-coupled dictionary learning SISR method was not better than Xu's method for all 14 images. For test images Barbara, Kodak-08, Nu-Regions and Peppers Xu's method would give better PSNR and for the remaining 10 images the proposed method would give higher PSNR values. Curious about why this was so, we calculated the power spectral density for all images in Set-A which is as shown in figure 5.1. We note that for images that had most of its frequency content in and around the low frequency region the proposed method would give better results and for images whose frequency content is wide and spread then Xu's method would perform better.

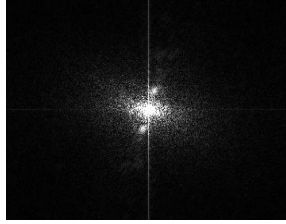
Table 5.5: Proposed SCDL SISR method versus classic bi-cubic interpolator, Yang's algorithm and Xu's algorithm.

Images	Bic.	Yang	Xu	Proposed
AnnieYukiTim	31.424	32.853	32.800	32.960
	0.906404	0.938103	0.9375	0.922934
Barbara	25.346	25.773	25.866	25.821
	0.792961	0.832917	0.8340	0.834083
Butterfly	27.456	30.387	30.047	30.638
	0.898450	0.948782	0.9447	0.938459
Child	34.686	35.420	35.405	35.420
	0.841002	0.880024	0.8797	0.864272
Flower	30.531	32.442	32.284	32.536
	0.896804	0.933642	0.9326	0.926663
HowMany	27.984	29.219	29.182	29.258
	0.868694	0.91325	0.9125	0.88598
Kodak-08	22.126	22.827	23.507	22.871
	0.699514	0.767805	0.7964	0.757098
Lena	35.182	36.889	36.851	37.000
	0.921749	0.948912	0.949	0.936711
MissionBay	26.679	28.012	27.929	28.115
	0.845938	0.890611	0.8883	0.880447
NuRegions	19.818	21.383	22.074	21.673
	0.846978	0.906466	0.9178	0.91118
Peppers	29.959	31.283	31.996	31.358
	0.904532	0.944194	0.9588	0.911413
Rocio	36.633	39.217	39.075	39.285
	0.961259	0.978078	0.9775	0.970202
Starfish	30.225	32.205	31.967	32.363
	0.892305	0.936731	0.9350	0.924364
Yan	26.962	28.022	28.003	28.107
	0.827686	0.874876	0.8743	0.856287
Average	28.929	30.424	30.499	30.529
	0.864591	0.906742	0.90986	0.894292

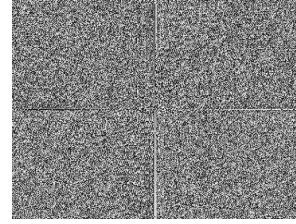
AnnieYukiTim.bmp



Magnitude of FFT2



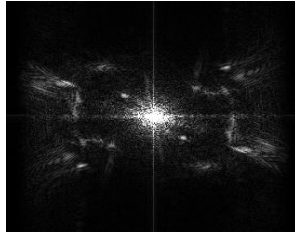
Phase of FFT2



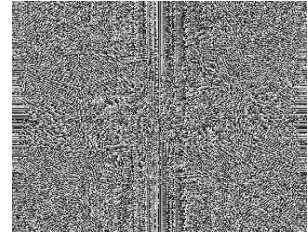
Barbara.png



Magnitude of FFT2



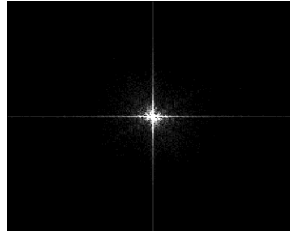
Phase of FFT2



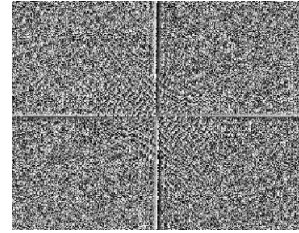
child.jpg



Magnitude of FFT2



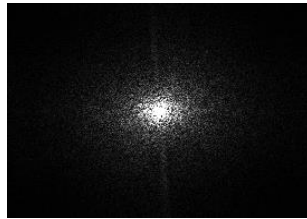
Phase of FFT2



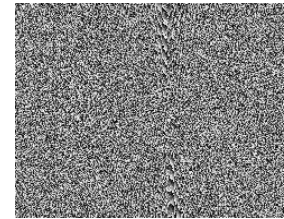
flower.jpg



Magnitude of FFT2



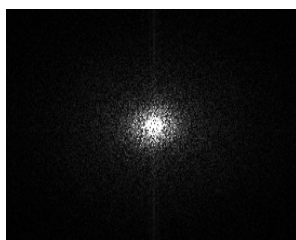
Phase of FFT2



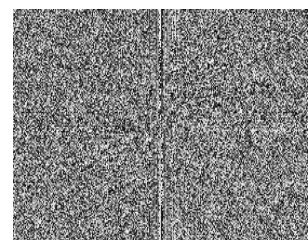
HowMany.bmp



Magnitude of FFT2



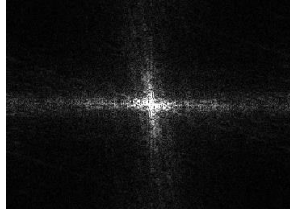
Phase of FFT2



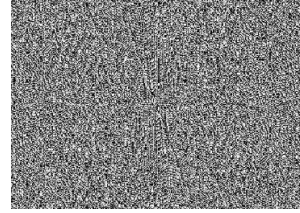
Kodak_8.bmp



Magnitude of FFT2



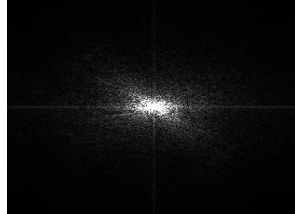
Phase of FFT2



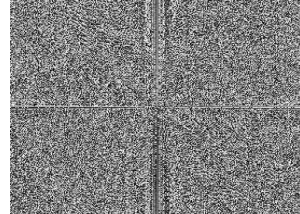
lena.jpg



Magnitude of FFT2



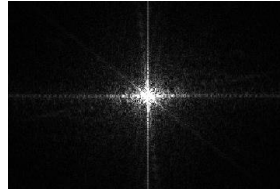
Phase of FFT2



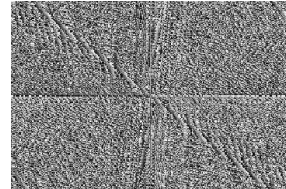
MissionBay.bmp



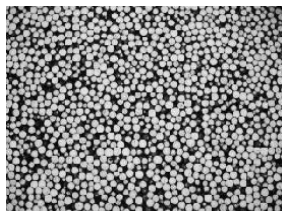
Magnitude of FFT2



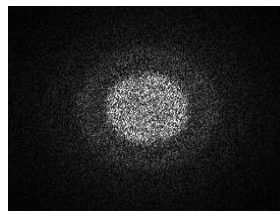
Phase of FFT2



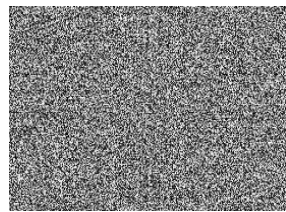
NuRegions.bmp



Magnitude of FFT2



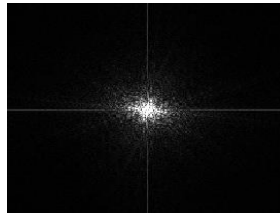
Phase of FFT2



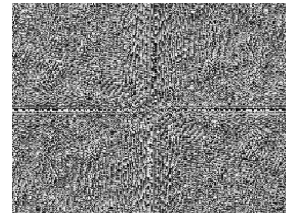
Peppers.bmp



Magnitude of FFT2



Phase of FFT2



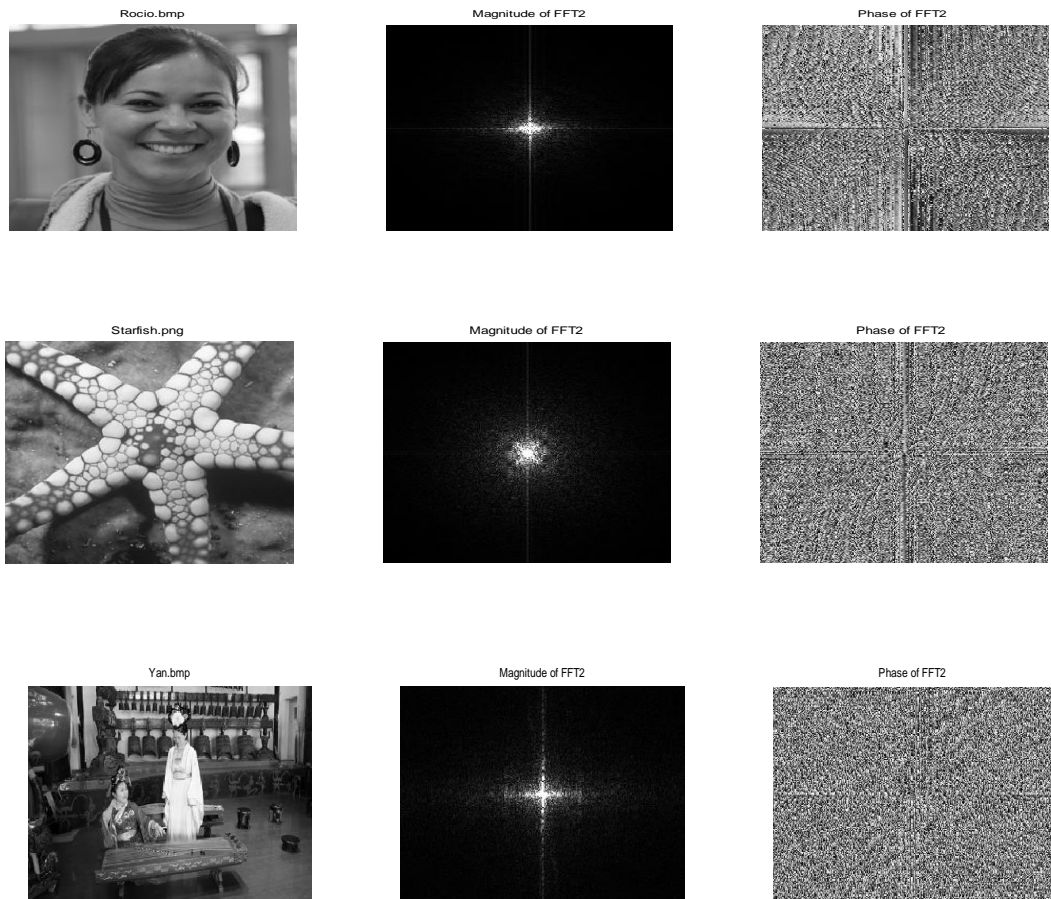


Figure 5.1: The Power Spectral Density Plots of Different images in Set-A

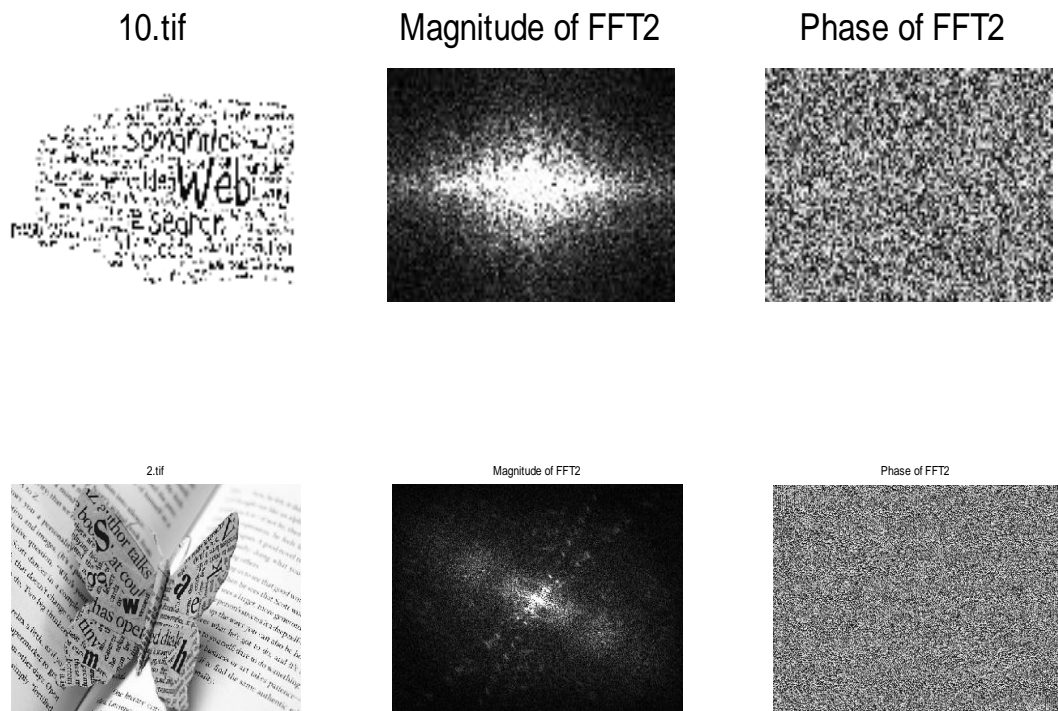
In the second set of simulations which used test images in Set-B, only the two best performing algorithms were compared. The results obtained are as depicted in Table

Table 5.6: Xu versus Proposed Method on test images in Set-B

Xu			Proposed SCDL SISR method		
Images	PSNR	SSIM	Images	PSNR	SSIM
10.tif	24.3158	0.9472	10.tif	24.5297	0.9561
2.tif	22.7973	0.8487	2.tif	23.0514	0.8503
5.tif	22.5205	0.907	5.tif	23.4412	0.9313
6.tif	25.5222	0.9478	6.tif	25.7378	0.9435
b82	27.3879	0.925	b82	27.4646	0.9202
t1	24.3933	0.8898	t1	22.2233	0.8345
t2	18.7391	0.7532	t2	16.3692	0.6655
t3	19.717	0.7741	t3	18.7558	0.7438
t4	21.22	0.6885	t4	19.2275	0.5746
Yxfo16	22.8599	0.8543	Yxfo16	23.0263	0.8482

In testing Set-B, images 10.tif, 2.tif, 5.tif, 6.tif, t1.jpg, t2.jpg, t3.jpg and t4.jpg are text images either in gray tones or in colour. Interestingly, for half of these text images the proposed algorithm will give higher PSNR values (marked in bold) and for the other half Xu's method would. In the mean PSNR values the proposed method has 0.1664dB edge over Xu's method. To see if our previous PSD based argument would hold for this set of images also, we again plotted the power spectral densities for test images in set-B which are as depicted in Table 5.6. To our surprise, even when the PSD contained many different high frequency components (spread psd plot) sometimes Xu and sometimes the proposed method would provide higher PSNR values. For test images t1.jpg, t2.jpg, t3.jpg and t4.jpg the proposed algorithm would give higher PSNRs and for 2.tif, 5.tif, 6.tif and 10.tif Xu's method would. Clearly a second factor must be tipping the balance to one or the other method. To better understand why this was so, we followed an idea on scale-invariance proposed in [1]. Three clusters are denoted by C1 through C3 corresponding to sharpness

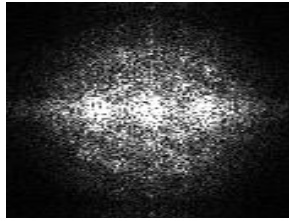
measure (SM) intervals of [0, 5], [5, 10], [10, 20] were defined. First, the sharpness measure values of all HR and LR patches of the image were calculated. Patches were then classified into the three clusters C1-C3 based on the calculated sharpness measure values and selected intervals. The number of total high-resolution patches divided into each interval was counted. Also, LR counterparts were appropriately classified into the same cluster were counted. Based on these counts, SM invariance was calculated as a ratio of LR patches correctly classified to the entire number of HR patches in a cluster. These SI-values are provided in Table 5.7.



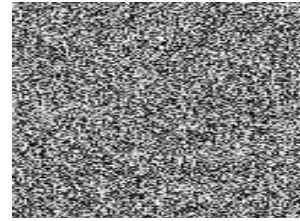
5.tif



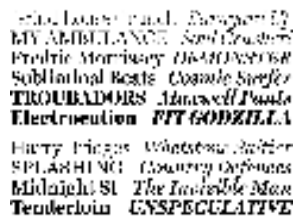
Magnitude of FFT2



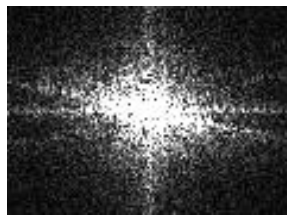
Phase of FFT2



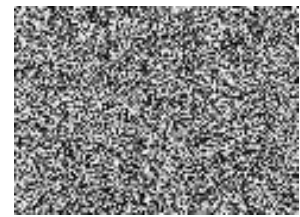
6.tif



Magnitude of FFT2



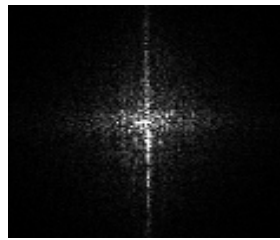
Phase of FFT2



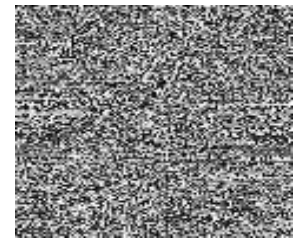
b82.tif



Magnitude of FFT2



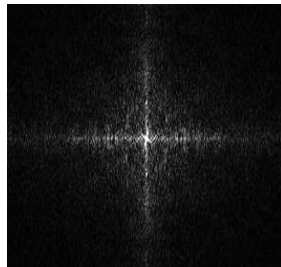
Phase of FFT2



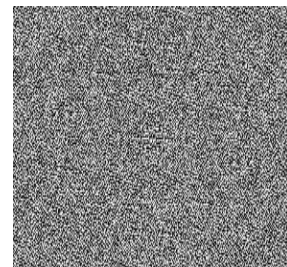
t1.jpg



Magnitude of FFT2



Phase of FFT2



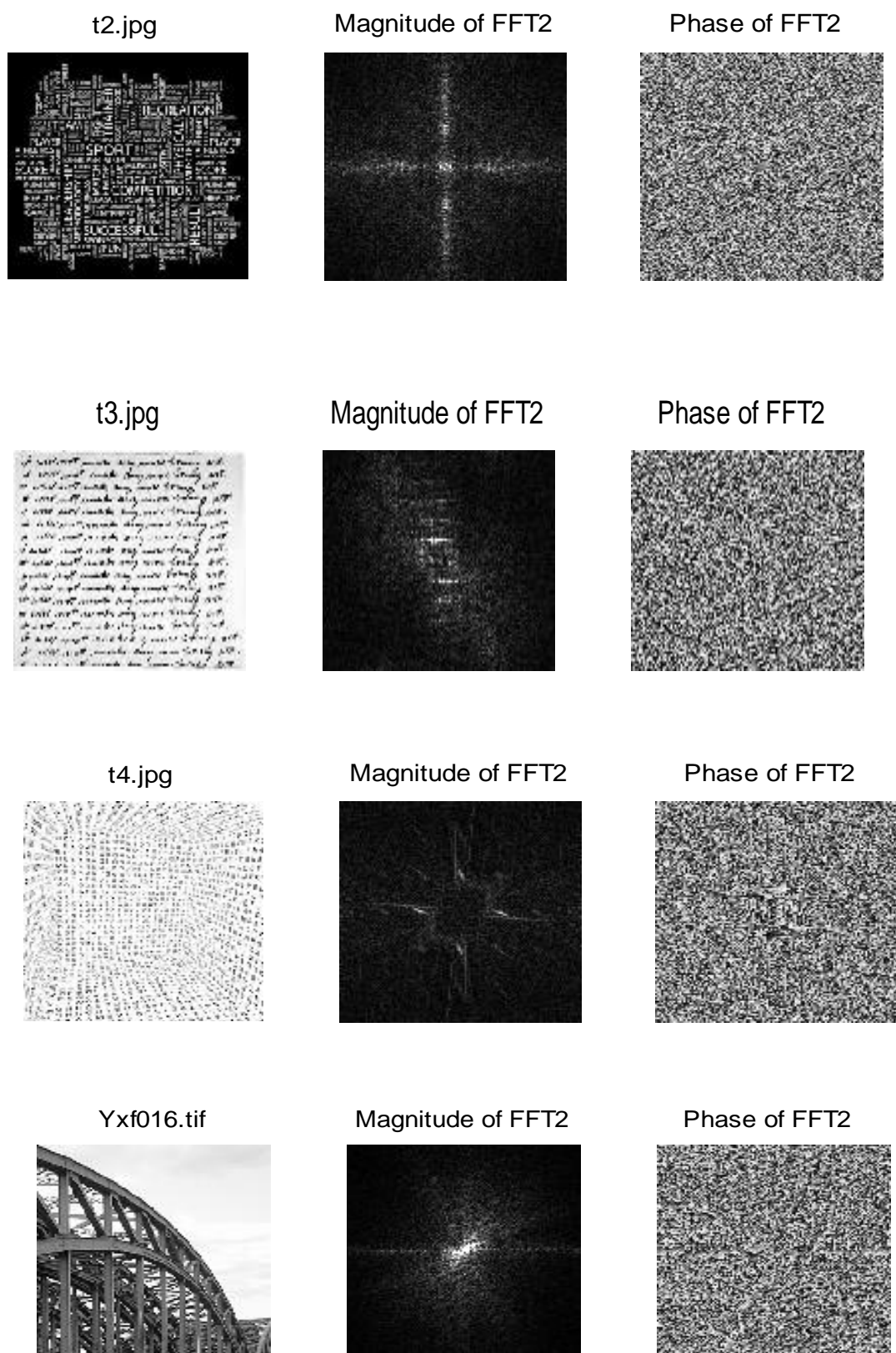


Figure 5.2: Power Spectral Density Plots of images in Set-B

Table 5.7 Scale Invariance of Regular Texture Images and Text images entire numbers of HR patches categorized in each interval (top) SM invariance (bottom)

Image	C1	C2	C3
10	31,864	936	11,570
	94.15328	38.3547	97.74417
2	3,173	898	2,529
	98.6133	50.33408	51.16647
5	10,502	672	7,426
	94.79147	75.44643	94.68085
6	26484	983	11,509
	91.69687	26.14446	99.18325
b82	463	700	1,438
	90.49676	77	57.30181
t1	960	71	949
	98.85417	64.78873	55.00527
t2	607	16	1,402
	99.17628	75	64.47932
t3	406	328	1,255
	92.11823	40.54878	75.21912
t4	36	118	1,871
	52.77778	61.86441	9.353287
yxf016	1,173	274	1,154
	97.78346	52.91971	53.37955

It was noted that for images with many frequency components (spread PSD) when the number of HR patches in C2 and/or C3 was low and their corresponding SM-invariance ratios were also low then the proposed method will not be as successful as Xu's method. For images that's PSD has frequency components mostly at low frequencies the proposed method would always perform better than Xu's method.

To further test this idea we selected 8 more images from the Kodak set that had wide PSDs and also calculated their corresponding SM-invariance values which are given in Table 5.8. For all cases the results were as expected, whenever for C2 and/or C3.

The number of HR patches classified in that interval plus the corresponding SM-invariance values were low, then as expected Xu’s algorithm will outperform the proposed method.

Table 5.8: Scale Invariance of Regular Texture Images and Text images entire numbers of HR patches categorized in each interval (top) SM invariance (bottom)

Image	C1	C2	C3
AnnieYukiTim	4985	1092	859
	99.29789	44.87179	47.6135
Barbara	5414	2012	2978
	99.64906	41.6501	19.37542
BooksCIMAT	1115	865	934
	98.74439	60.34682	53.31906
Fence	1068	598	935
	99.1573	27.59197	37.75401
ForbiddenCity	710	580	1624
	98.87324	48.7931	4.248768
Michoacan	676	372	1494
	37.57396	36.29032	30.25435
NuRegions	46	112	2384
	4.347826	29.46429	97.86074
Peppers	1481	371	357
	98.64956	63.8814	48.7395

5.3 Quality of Super-Resolved Images

In figures 5.3-5.7 we have provided the SR images produced by the competing methods for visual comparison. Images are zoomed to clarify the reconstruction quality. In line with the PSNR values, the bi-cubic images show a significant amount of blur however the other images show comparatively less blur. Looking at the zoomed images it can be clearly observed that’s proposed method is on par with the state of the art algorithms and is able to recover the sharp patches effectively.

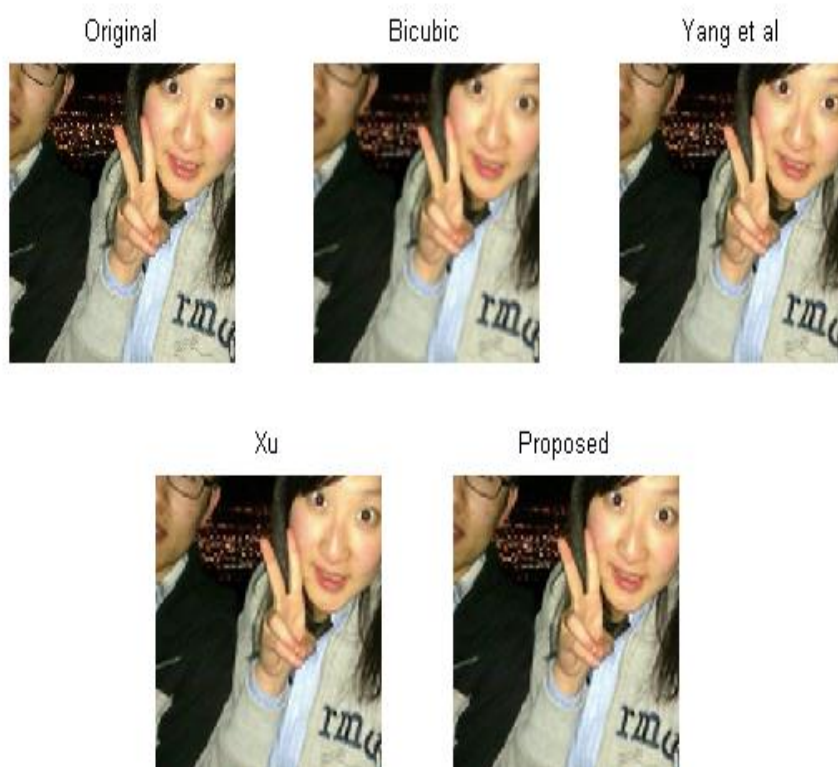


Figure 5.3: SR images for test image **AnnieYukiTim** from Set-A.



Figure 5.4: SR images for test image **Flower** from Set-A



Figure 5.5: SR images for test image **Butterfly** from Set-A.

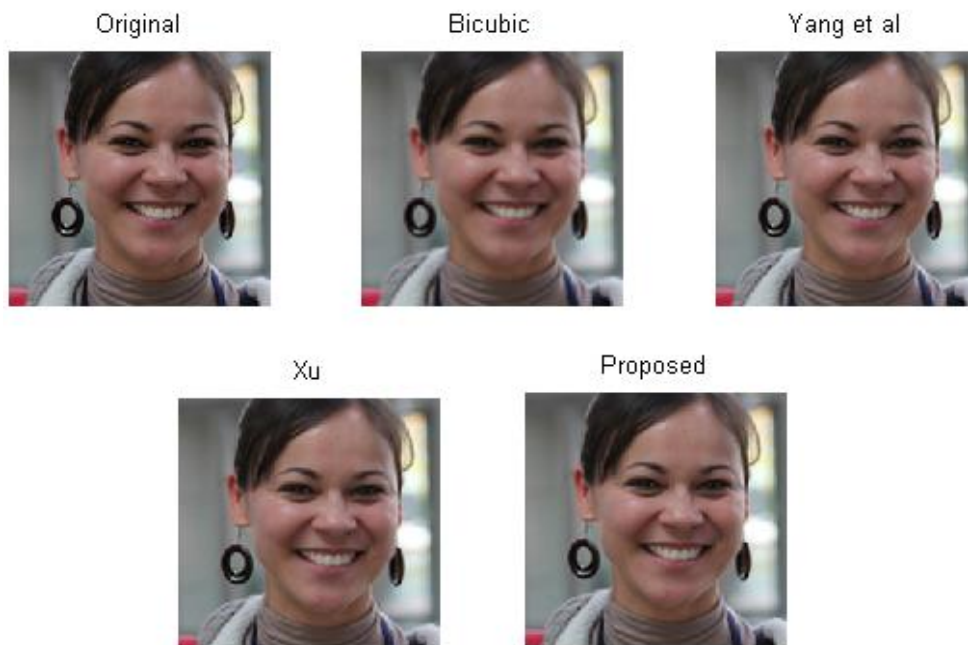


Figure 5.6: SR images for test image **Rocio** from Set-A.

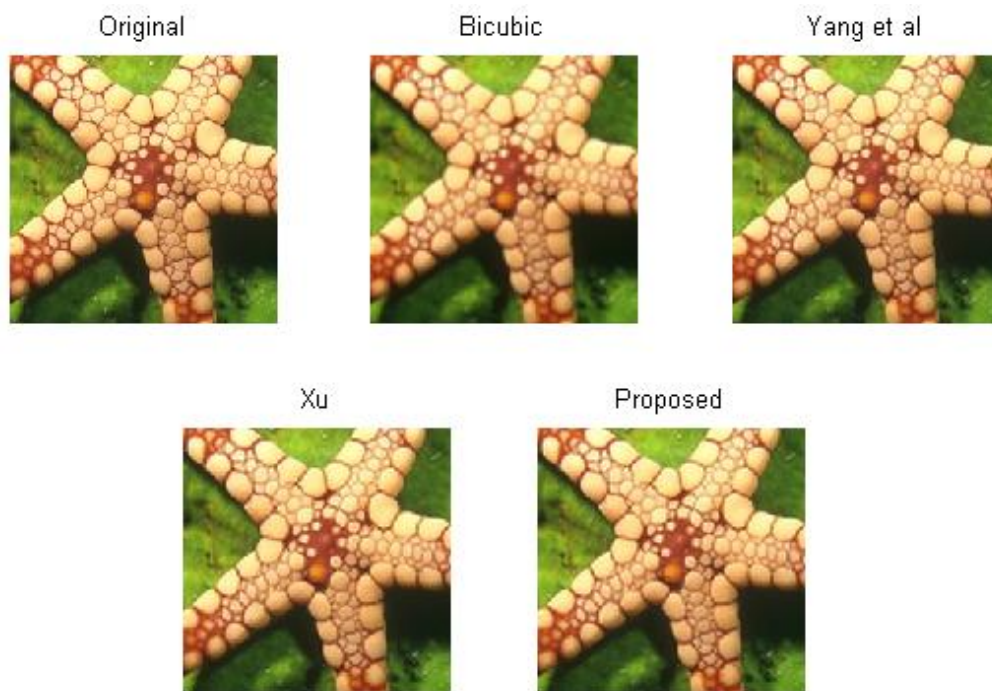


Figure 5.7: SR images for test image **Starfish** from Set-A

Chapter 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In our work we proposed a semi-coupled dictionary learning strategy for single image super-resolution. An approximate scale invariant feature due to resolution blur is proposed. The proposed feature is used for classifying the image patches into different clusters. It is further suggested that the use of semi-coupled dictionary learning, with mapping function that can recover the representation quality. The proposed strategy for SISR contains on two phases, the first phase is about the training and the other one is the reconstruction phase. In the training phase a set of HR images are taken as input, and these HR images are blurred down-sampled for the purpose to get Low resolution images. Bi-cubic interpolation is done on these LR images to get MR images, and from these MR and HR images the patches are extracted and put in their respective cluster on its sharpness measure values to get HR and LR data training matrices. After this we initialize the dictionaries D_H and D_L and also initialize the mapping functions T_H and T_L and in last updating stage applied on all these dictionaries and get the updated dictionaries. In reconstruction phase an LR test image with pair of dictionary with mapping function are taken and with the help of this the patches are recovered from each cluster and find the sharpness value for each LR patch and putted in their expected clusters. From LR dictionary we find the LR patches and from HR dictionary find the HR patches using sparse invariance property. In last we reshape and merge the HR patches to recover

the approximate high-resolution image. The comparison of proposed method is made with bi-cubic, Yang and Xu, and for this purpose two sets of images namely Set-A and Set-B were used in which it showed that which algorithm gives high PSNR values. In Set-A 14 images were used in which the proposed method gives better PSNR values from bi-cubic and also from the method of Yang and Xu. In Set-A images the conflict come between the proposed and Xu methods, and it was observed from the results of some images. In which the results of Xu is better for some images and for some the proposed method performs better, to clarify this doubt a PSD term is used in which it shows that for those images in which the frequency contents at the low frequency region and PSD plot is not spread in this case the proposed algorithm will perform better. In the images in which the frequency contents spread out, the performance of Xu will be better. In Set-B the images are text images and its gives some surprising results, in which for half images the Xu method perform better and for other half the proposed methods perform better. To solve this problem that why the images behaviour are like this, for this purpose scale invariance technique is used which gives an evident about this problem.

In the proposed algorithm a patch size of 5×5 with 600 dictionary atoms is used and comparison is made with the ever green Bi-cubic, Yang and Xu. The performance of the proposed algorithm illustrate 1.6 dB improvement and SSIM 0.02 over bi-cubic interpolation while also ahead from Yang and Xu in term of PSNR 0.10 dB and 0.03 dB and inferior in SSIM results 0.012 and 0.012 respectively.

6.2 Future Work

In this thesis we proposed a semi-coupled dictionary learning algorithm for single image super-resolution, which contain number of dictionaries with a mapping function and number of clusters to obtain a high resolution image. This work can be extended to increase the number of dictionaries as well as the number of clusters and it would be interesting to optimize the SM intervals in a more systematic manner. For example bounds of intervals perhaps can be selected to optimize the performance of the proposed algorithm.

REFERENCES

- [1] Yeganli, F., Nazzal, M., Ünal, M., and Özkaramanlı, H. (2015). Image super resolution via sparse representation over multiple learned dictionaries based on edge sharpness, *J of Signal Image and Video Processing*, (pp. 535–542).
- [2] Yang, J., et al. (2010). Image super-resolution via sparse representation, *IEEE Trans. on Image Processing*, (pp. 2861-2873).
- [3] Yang, J., Wang, Z., Lin, Z., Cohen, S., Huang, T. (2012). Coupled dictionary training for image super-resolution, *IEEE Trans. on Image Processing*, (pp. 467-3478).
- [4] Xu, J., Chun, Q., & Zhiguo C. (2014). Coupled K-SVD dictionary training for super-resolution, *Image Processing (ICIP), IEEE Conference*, (pp. 3910-3914).
- [5] 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*, (pp. 2216-2223).
- [6] Zhang, K., et al. (2012). Multi-scale dictionary for single image super-resolution, *Computer Vision and Pattern Recognition (CVPR), IEEE Conference*, (pp. 1114-1121).

- [7] Jia, K., Xiaogang W., & Xiaoou T. (2016). Image transformation based on learning dictionaries across image spaces, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (pp. 367-380).
- [8] 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.
- [9] Wang, Z., Bovik, AC., Sheikh, HR., Simoncelli, EP. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on Image Processing*, (pp. 600-612).
- [10] Liu, D., Klette, R. (2015). Sharpness and contrast measures on videos, In: *Proceedings of International Conference on Image Vision Computing (IVCNZ'15), IEEE Conference*, (pp. 1-6).
- [11] Bruckstein, A.M., Donoho D.L., and Elad, M. (2009). From Sparse solutions of systems of equations to sparse modeling of signals and images, *J of SIAM Review*, (pp.34-81).
- [12] Kodak lossless true color image suite. (20 January 2016).
online:r0k.us/graphics/kodak/index.html.

- [13] Tsai, R.Y., Huang, T.S. (1984). Multi frame image restoration and registration. In: *Advances in Computer Vision and Image Processing*, (pp. 317 – 339).
- [14] Nguyen, N., Milanfar, P. (2000). A wavelet-based interpolation-restoration method for superresolution, *Circuits Syst. and Signal Process.* (pp. 321–338).
- [15] Chappalli, M.B., Bose, N.K. (2005). Simultaneous noise filtering and super-resolution with second-generation wavelets. *IEEE Signal Process.*,(pp. 772775).
- [16] Ur, H., Gross, D. (1992). Improved resolution from subpixel shifted pictures, *Graphical Models and Image Processing*, (pp. 181–186).
- [17] Patti, A.J., Tekalp, A.M. (1997). Super resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Trans. on Image Process.*, (pp.1446–1451).
- [18] Chang, H., Yeung, D.Y., Xiong, Y. (2004). Super-resolution through neighbour embedding. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, (pp. 275–282).
- [19] Zibettia, M.V.W., Baznb, F.S.V., Mayer, J. (2011). Estimation of the parameters in regularized simultaneous super-resolution, *J of Pattern Recognit.* (pp. 69–78).

- [20] Tom, B.C., Katsaggelos, A.K. (1995). Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. In: *Proceedings of the IEEE International Conference on Image Processing*, (pp. 539–542).
- [21] Woods, N.A., Galatsanos, N.P., Katsaggelos, A.K. (2006). Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *IEEE Trans. on Image Process.*, (pp. 201–213).
- [22] Dempster, P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J of the Royal Statist. Soc.*, (pp.1–38).
- [23] Bertero, M., Boccacci, P. (1998). *Introduction to Inverse Problems in Imaging*. IOP Publishing Ltd, Philadelphia.
- [24] Hardie, R. C., Barnard, K.J., Armstrong, E.E. (1997). Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. on Image Process.*, (pp. 1621–1633).
- [25] Schultz, R.R., Stevenson, R.L. (1996). Extraction of high-resolution frames from video sequences. *IEEE Trans. on Image Process.*, (pp. 996–1011).
- [26] Hertzmann, A., Jacobs, C.E., Oliver, N., Curless, B., Salesin, D.H. (2001), Image analogies, In: *Proceedings of the SIGGRAPH*, (pp. 327–340).

- [27] Chang, H., Yeung, D. Y., Xiong, Y. (2004). Super-resolution through neighbor embedding. In: *Proceedings of the IEEE International on Computer Vision and Pattern Recognition*, (pp. 275–282).
- [28] Tian, J., Ma, K-K. (2011). A survey on super-resolution imaging, *J of Signal Image and Video Processing*, Vol.5, Issue 3, (pp. 329-342).
- [29] Donoho, D. L. (1995). “De-noising by soft-thresholding” *IEEE Trans. on Inform. Theory*, Vol. 4, No. 3, (pp.613-627).
- [30] Engan, K., Aase, S.O., and Husoy, J.H. (1999). Method of optimal direction for frame desing, In: *Proc. of Acoust. Speech and Signal Process.*, (pp. 2443-2446).
- [31] Aharon, M., Elad, M., and Bruckstein, A. (2006). “K-svd: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. on Image Process.*, vol.54, no, 11 (pp. 4311-4322).
- [32] Lee, H., Battle, A., Raina, R., and Ng, A. Y. (2007). Efficient sparse coding algorithms, *Advances in Neural Information Process. Sys.*, Vol. 19, No. 2, (pp. 801-808).
- [33] Mairal, J., Elad, M., and Sapiro, G. (2008). Sparse representation for color image restoration, *IEEE Trans. on Image Process.*, Vol. 17, No. 1, (pp.53-69).

- [34] Elad, M., Starck, J. L., Querre, P., and Donoho, D. L. (2005). Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), *J of Applied and Computational Harmonic Analysis*, Vol. 19, (pp. 340-358).
- [36] Elad, M., and Aharon, M., (2006). Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Transactions on Image Processing*, Vol. 15, (pp. 3736-3745).
- [37] Glasder, D., Bagon, S., and Irani, N., (2009). Super-Resolution from Single Image, *IEEE 12th International Conference on Computer Vision*, (pp. 349-356).
- [38] MacQueen, J., (1967). Some methods for classification and analysis of multivariate observations, *Proceedings of the Fifth Berkeley Symposium On Mathematical Statistics and Probabilities*, Vol.1, (pp. 281-296).