Chapter 1

Introduction

1.1 Image super-resolution

The technique of obtaining a high-resolution (HR) image from corresponding single or multiple low-resolution (LR) image(s) along with enhancements of the underlying detail features is known as super-resolution (SR) imaging. Spatial resolution of an image is calculated as the number of pixels used to represent the image, where each pixel embodies the smallest distinguishable element of the image. While, resolution of a digital camera is specified by the spatial density of the optical sensor; low-cost sensors generally have less spatial density. If an image is captured by a LR sensor (e.g. QVGA: 320×240) and displayed using a HDTV (1024×720) display, it will be distorted and blurred. Objects in the image will become unrecognizable. This is due to the fact that the number of pixels available in the LR image are stretched to fit the HDTV screen's dimensions. To overcome this, SR imaging employs signal processing techniques to recover the missing pixels in a given LR image in order to obtain the corresponding HR image. The imaging method also includes steps for removal of blur effects induced during the capture of the input LR image(s).

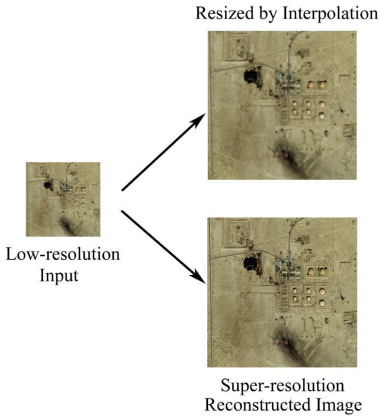


Figure 1.1: Image super-resolution vs. interpolation

Fig. 1.1 shows twofold enhancement of resolution of an image using SR imaging. SR reconstructed image has clear evidence of containing more sharp details than the image resized via simple nearest neighbour interpolation (NNI) technique. Resolution enhancement of images has urged very much importance now-a-days owing to the rapidly growing high-definition (HD) applications in the areas of video surveillance, medical diagnosis, remote sensing, and many more [2, 19, 40, 74, 115, 131].

1.2 Super-resolution in remote sensing

Environmental monitoring, disaster management, land-cover mapping, and other remote sensing applications necessitate images with high spatial/spectral resolutions. The high spatial and spectral resolutions of remote sensing images can provide more precise geometric and thematic studies. Despite the fact that multispectral (MS) images typically contain significant amount of spectral information, their spatial resolutions are poor. HR MS images are used in a variety of remote sensing applications, where LR images fail to deliver high-quality detail information for meaningful analysis.

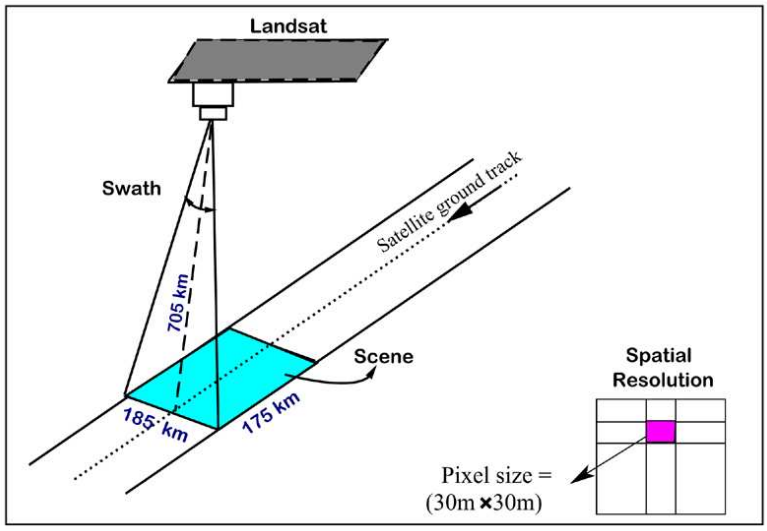


Figure 1.2: Demonstration of image acquisition by Landsat satellite

SR is very important for remote sensing applications as most satellites capture images at relatively lower spatial resolutions. Fig. 1.2 shows a schematic of the

image acquisition process by a satellite. Since these images cover a large area (say, 50-100 square kilometers) on the earth, an object in such an image is hardly recognizable. Besides, images are also degraded due to the satellite movements, atmospheric condition, etc. Two MS images with spatial resolutions of 30 m (LR) and 10 m (HR) acquired from the same location are shown in Fig. 1.3.

Figure 1.3: Example of satellite image super-resolution

Remote sensing satellites once lunched into the orbit, it is not possible to replace their integrated LR imaging sensors with the recent one. As, the ground stations keep receiving such LR images regularly, a SR software module installed in the base station can be very handy for processing images from these LR satellites. It also helps in reduction of costs and time; promoting HR-based applications using existing LR satellites. Moreover, a carefully developed embedded hardware module capable of performing SR can be placed on-board in an aircraft for obtaining HR imagery.

1.2.1 Limitations of remote sensing image SR

Satellite images are generally very big (e.g. 7000×8000 pixels). Moreover, multispectral images consist of several bands. So, the *data size* becomes an issue for remote sensing image SR. Additionally, it is a challenge to estimate the corresponding *spectral features*, while targeting the HR spatial reconstruction. Therefore, the *computational complexity* of MS SR methods becomes high as it needs to upscale all the spectral bands at a time. Nevertheless, creating a standard database is an issue as the images need complicated pre-processing steps, like radiometric and geometric corrections. For example, there are no publicly available *database* for the Indian multispectral satellite ResourceSat-2.

1.3 SR approaches and their limitations

One straightforward way to get images with higher spatial-resolution is to upgrade the imaging system by replacing the sensor with technological advancements. There are two possible ways for this as follows [76]:

(i) *Reduce the pixel size:*
It is possible to acquire images with a higher pixel count per unit area using cutting-edge sensing technologies. However, this results in smaller pixel sizes, and as a result, less average incident light on each pixel. As a result, more shot noise is generated, degrading the image quality even further.

(ii) *Increase the chip size:*
By increasing the chip size, it allows to accommodate more number of transistors on the wafer resulting in high spatial density. However, as the chip size is increased the capacitance is also increased resulting in a slow rate of charge transfer.

Nonetheless, the high cost of latest HR sensors is also a matter of concern. In such scenarios, generation of HR image through super-resolution imaging is more beneficial as it allows to use the existing low-cost LR sensors for remote sensing applications [6, 74].

1.4 Fundamentals of single image super-resolution

Super-resolution (SR), in general, is defined as an inverse problem in digital image processing. Single image SR (SISR) aims at recovering the HR version of a given LR observation. The LR image is viewed as the degraded version of the HR image due to many reasons, like blurring, downsampling, and some additive noise due to measurement error. Therefore, SISR makes an estimate of the unknown HR

image by reversing the image degradation process. In other words, SR is the inverse mapping of a degradation process.

1.4.1 Image acquisition/degradation model

SR image reconstruction can be formulated from the basic image degradation model. An observed image is actually a blurred and downsampled approximation of a real-world scene. The camera len's finite aperture leads to diffraction; the point spread function (PSF) varies spatially and usually results in the blurring effect. Additionally, due to the integration of limited transistors, a digital sensor can capture only a selective number of sample pixels. It causes downsampling of the original HR scene.

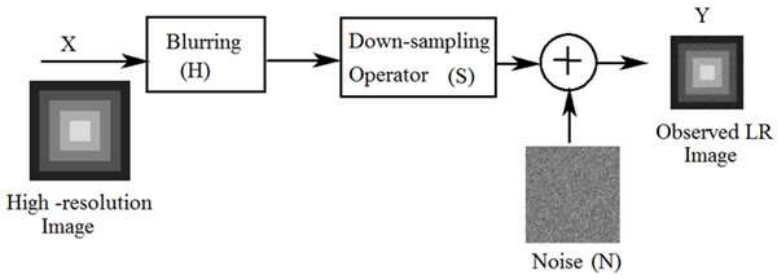


Figure 1.4: Image degradation model

Fig. 1.4 shows the basic image degradation model where the observed LR image $\mathbf{Y}$ is a blurred and downsampled version of the original HR image $\mathbf{X}$. Mathematically, we the representation can be written as,

$$\mathbf{Y} = SH\mathbf{X} + N, \tag{1.1}$$

where H represents a blurring operator, and S is downsampling operator. For super-resolution problems, the additional noise parameter N is considered ideally equal to zero or negligible. Recovering the original HR image is treated as an ill-posed inverse problem as for any LR input $\mathbf{Y}$ infinitely many HR images $\mathbf{X}$ may satisfy the above equation. Also, the blurring operator generally remains uncertain for different imaging conditions. Moreover, the multiplication SH results in an overcomplete rectangular matrix, thus the system represented by Eq. 1.1 becomes an underdetermined linear system [76].

1.4.2 Image degradation

For simulation of SR imaging, generally the available image is assumed to be the given HR image. Then, a corresponding LR image is synthesized from it by applying the image degradation process. The image degradation model is given by:

$$\mathbf{I}_{LR} = S\left(H\left(\mathbf{I}_{HR}\right)\right) + N \tag{1.2}$$

Degradation is the reverse process of SR where we first blur the HR image using a low pass filter followed by downscaling to get the subsampled LR image [41]. Both the PSF for blurring operator H, and additive white Gaussian noise N can be adequately approximated by the Gaussian distribution function:

$$G\left(x\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{1.3}$$

where the parameters μ and σ denotes mean and standard deviations, respectively. Gaussian filtering is carried out by taking a mask or kernel of size (denoted by k) 5×5 or 7×7 or higher and convolves it with the image by shifting the mask starting from top-left. The degree of blur depends on the value of σ, a higher value of σ supported with a larger mask will cause more blur. Ideally, the Gaussian mean μ is taken as zero. Fig. 1.5 presents a visual of blurred and downsampled image from a test image using a 5×5 mask with $\sigma = 1.2$ and then downscaling by two.

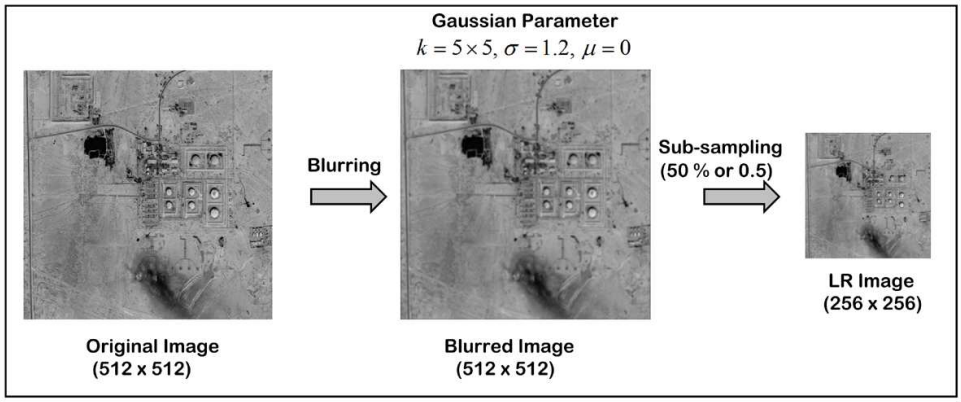


Figure 1.5: Process of LR image synthesis from an image

1.4.3 Super-resolution versus interpolation

Although some literature refer image interpolation as a SR technique [60], there is some difference in their working. Fig. 1.6 shows an example of image interpolation, where the LR image is mapped into a target HR grid and the missing pixels in the new grid are estimated using interpolation functions, like bilinear, bicubic, etc. They simply do not consider any process to restore the high-frequency information in target HR image, albeit a few interpolation methods are also capable of slightly reducing the aliasing effects caused by the downsampling. On the contrary, SR methods are more dedicated techniques making efforts to improve the resolution as well as restoration of HR information as much as possible.

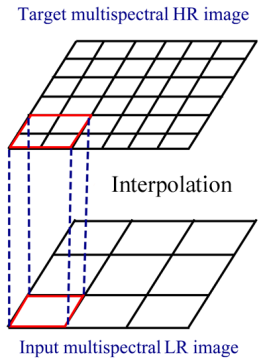


Figure 1.6: Interpolation of pixels for upscaled image

1.4.4 Super-resolution versus pansharpening

In remote sensing, pansharpening are the methods that combine an HR panchromatic image with a LR MS image to obtain the target HR MS image. Although, both the SR and pansharpening techniques result into an HR MS image, their distinguishing fact is that the latter targets to restore the missing spectral information for the panchromatic (PAN) band, whereas the SR method tries to restore both the spatial and spectral information, simultaneously from the given LR MS image.

1.4.5 Approaches for SR of MS images

To obtain SR of a colour image, first RGB to YC_bC_r transformation is done where the luminance channel Y contains the most of the high-frequency information. The other two channels, C_b and C_r, i.e. the blue- and red-difference chroma components give the colour information. Figure 1.7 depicts the extracted Y, C_b and C_r channels from a given RGB satellite image. SR is performed on the Y-channel only and the remaining two channels are simply upscaled via an interpolation technique. Finally, YC_bC_r to RGB inverse transformation is used to achieve the the target SR reconstructed RGB image.

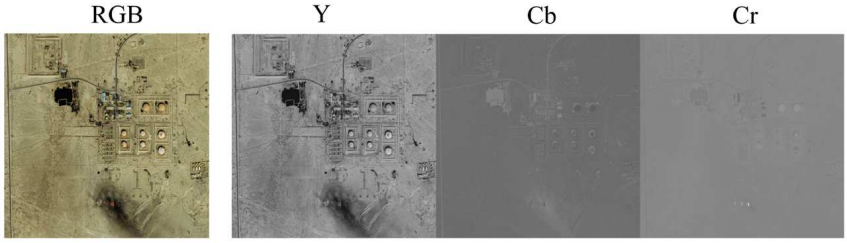


Figure 1.7: Visualization of Y, C_b, C_r bands obtained from an RGB image

We know that MS images consists of several bands (3–10) and one straightforward approach is to obtain a false color RGB image from the available bands (or the significant bands only) and apply the SR algorithm to the luminance channel. However, we observe that, by doing so the spectral properties of a MS image is causing spectral information loss. Therefore, SR is performed on each band separately of the given LR MS image to keep their individual spectral properties intact.

1.4.6 Evaluation parameters

In SR works, the reconstructed images are compared with the original image using different evaluation metrics for validation of the results. For simulation purpose, the ground-truth image is the given image (target HR) from which the test LR image is formed and reference-based evaluation parameters are computed based on them. Otherwise, if ground-truth is not available, no-reference-based parameters

are preferred and computed from the test LR image.

The peak signal-to-noise ratio (PSNR) expressed in dB and the mean structural similarity (MSSIM) index are two of the most often used reference-based metrics for objective evaluation. PSNR is inversely proportional to the mean-squared error between the ground-truth and the reconstructed images. The higher the PSNR, the better the reconstruction quality. On the other hand MSSIM measures the structural similarity between the two images whose value lies in the range [-1, 1]. For two exactly similar images, MSSIM value is 1. Some other well-known quantitative metrics used in objective evaluation, especially for super-resolution of MS images are:

(i) *Spatial Correlation Coefficient* (sCC) [78]: it measures linear relationship between edges of the reference image to that of the reconstructed image. It is expressed as:

$$sCC = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \tag{1.4}$$

where σ_{xy} denotes covariance between the ground-truth ($\mathbf{x}$) and the SR reconstructed image ($\mathbf{y}$). Similarly, σ_x and σ_y represent standard deviations of $\mathbf{x}$ and $\mathbf{y}$, respectively. Its value lies in the range [-1, 1].

(ii) *Universal Image Quality Index* (UIQI) [109]: it combines three different properties of image evaluation, namely, correlation, luminance and contrast. It is defined by:

$$Q = \frac{1}{K} \sum_{j}^{K} \left(\frac{\sigma_{xy}}{\sigma_x \sigma_y} \frac{2\mu_x \mu_y}{\mu_x^2 \mu_y^2} \frac{2\sigma_x \sigma_y}{\sigma_x^2 \sigma_y^2} \right)_j, \tag{1.5}$$

where μ_x, μ_y denotes the mean values of $\mathbf{x}$ and $\mathbf{y}$; and K represents the total number of MS bands. Its value also lies in the range [-1, 1].

(iii) *Erreur Relative Globale Adimensionnelle de Synthese*(ERGAS) [84]: it considers scaling factor and RMSE values for quality evaluation and expressed as follows:

$$ERGAS = \frac{100}{S} \sqrt{\sum_{k}^{n} \left(\frac{RMSE\left(\mathbf{x}, \mathbf{y}\right)}{\mu_x} \right)^2}, \tag{1.6}$$

where S is the SR scaling factor. ERGAS value of 0 indicates the best quality and higher values indicate distortions in the reconstructed output.

(iv) *Spectral Angle Mapper* (SAM) [33]: it is widely used for spectral assessment of MS images, finds the average angle between $\mathbf{x}$ and $\mathbf{y}$ assuming each band as a coordinate axis.

$$SAM = \frac{1}{P}\sum_{i}^{P} \arccos \frac{\mathbf{x}_i.\mathbf{y}_i}{\|\mathbf{x}_i\| \, \|\mathbf{y}_i\|}, \tag{1.7}$$

where P denotes the image's entire pixel count. Ideally, SAM should be 0.

(v) *Natural Image Quality Evaluator* (NIQE) [67]: it is a metric for image quality with no-reference; assumes that a distortion free image can be statistically represented by a multivariate Gaussian (MVG) distribution. First, it considers fitting a MVG distribution approximately on a training set of distortion free images. Next, it calculates the distance between the test image features (e.g. normalized luminance value fitted to Gaussian distributions, etc.) to those fitted to previously learnt MVG distribution. A smaller value of NIQE indicates better reconstruction ability of the SR method.

1.5 Sparse representation problem

A natural image is *compressible* because when it is transformed to the frequency domain using an image transformation technique, like, the discrete cosine transform (DCT) or the discrete Fourier transform (DFT), which uses sinusoidal basis functions, only a few coefficients are found significant and the rest are either small or negligible. The non-significant transform coefficients may be discarded by magnitude thresholding, while retaining only the significant ones for representing the image. On the other hand, a signal is said to be *sparse* if it can be expressed or approximated as a linear combination of a few elementary signals or basis functions selected from a proper dictionary or transform. The basis functions or vectors are also called the 'atoms' of the dictionary [37, 115]. A natural image is not sparse in

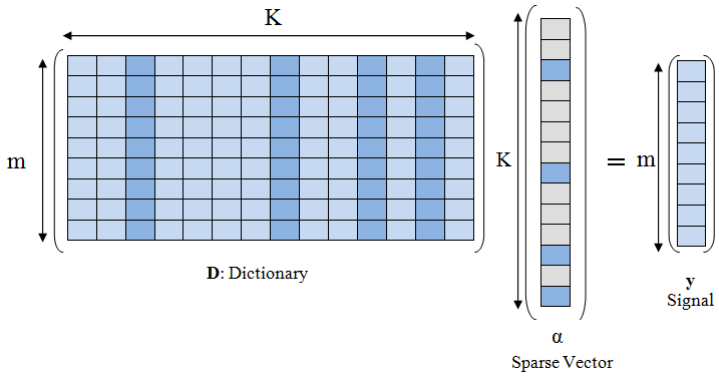


Figure 1.8: A sparse representation problem

any transform domain either by a fixed transform (e.g. DFT, DCT, etc.) or learned dictionary, but we can approximate it as a k-sparse signal having only k nonzero coefficients in the transform domain. *A sparse representation problem typically seeks a sparse solution using a learned overcomplete dictionary, where number of atoms sufficiently exceeds the dimension of the signal of interest* [14].

A pictorial demonstration of a typical sparse representation problem is given in Fig. 1.8. Here, we consider the problem of finding sparse vector $\boldsymbol{\alpha}_{K\times 1}$ such that $\mathbf{D}\boldsymbol{\alpha} = \mathbf{y}$, where vector $\mathbf{y}_{m\times 1}$ and matrix $\mathbf{D}_{m\times K}$ are given with $m \ll K$. It is an under-determined system with infinite candidate solutions but we want a sparse solution out of it. Therefore, in order to restrict the solution space, the problem can be reformulated with the inclusion of some *a priori* information about our signal of interest, i.e. the signal in question has a sparse representation. It follows:

$$\min_{\boldsymbol{\alpha}} \ \|\boldsymbol{\alpha}\|_0 \ \text{subject to} \quad \mathbf{D}\boldsymbol{\alpha} = \mathbf{y}, \tag{1.8}$$

where $\|\boldsymbol{\alpha}\|_0$ represents the ℓ_0-norm and gives the count of non-zeros present in $\boldsymbol{\alpha}$. Solving Eq. 1.8 is a non-deterministic polynomial time (NP) hard. Such ℓ_0-norm problems can be solved by either greedy algorithms or convex ℓ_1-norm minimization techniques [81]. A greedy algorithm, as the name implies, always selects the solution that appears to be the best at that time. This means that it makes a locally optimal decision in the hope of finding a globally optimal solution. However, they often fail to identify a globally optimal solution because they do not run through all of the data exhaustively. They may make hasty commitments to certain solutions,

preventing them from subsequently finding the best possible solution. Although these techniques are simple to use and computationally efficient, they do not provide recovery assurances, such as how well each sample can be reconstructed using the dictionary and sparse codes [13]. A method called 'Basis Pursuit' (BP) replaces ℓ_0-norm with ℓ_1-norm and thereby converting the non-convex problem into a convex optimization problem, i.e.

$$\min_{\boldsymbol{\alpha}} \ \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \ \ \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{y}\|_2^2 \leq \varepsilon \tag{1.9}$$

The above problem is the standard basis pursuit denoising (BPDN) problem [36]. It will find a sparse solution or even the sparsest one under certain conditions [86]. It is a convex optimization problem with accurate solvers like interior point methods already available [79]. A specific class of algorithms that are faster than interior point methods have been reported for larger problems. Algorithms in this category are based on soft-thresholding/shrinkage, such as the iterative shrinkage-thresholding (IST) [29] and fast iterative shrinkage-thresholding algorithm (FISTA) [8] to name a few.

1.6 Sparse representation in image processing

Sparse representation has attracted a lot of attention since it works well in a wide range of image analysis applications, including image inpainting, denoising, compression, demosaicking, and so on [36, 63, 88, 124]. Many researchers use sparse representation to build sparse dictionaries or to find sparse coefficients in existing dictionaries. The sparse representation approach, as shown in Fig. 1.9, tries to compute the sparse decomposition of each patch in the image by finding a linear combination of a few atoms/ bases from the learned dictionary [89].

If we assume $\mathbf{x}$ as our test patch, then it satisfies the following representation:

$$\mathbf{x} \approx \sum_{i=1}^{n} \boldsymbol{\alpha}_i \mathbf{d}_i, \tag{1.10}$$

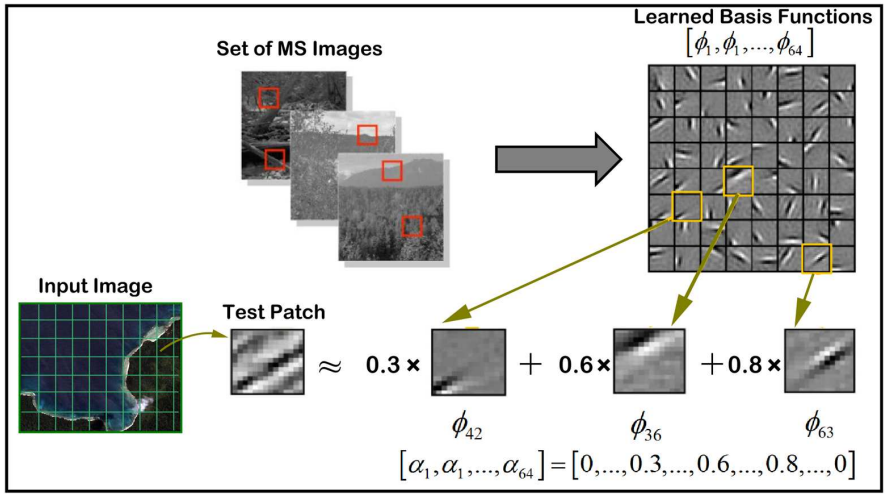


Figure 1.9: Example of sparse coding for an image patch using learned basis functions

where $\mathbf{d}_i$ represents basis functions/atoms. Again, n represents total number of atoms (e.g. n=64) and α represents the sparse coefficients vector (e.g. $\boldsymbol{\alpha} = [0, ..., 0.3, ..., 0.6, ..., 0.8, ..., 0]$) used to represent the patch. The dictionary atoms are trained from a corpus of similar images so that it is capable of representing the patches of input test image reasonably well. Assuming $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_n]$ be the collection of atoms forming an overcomplete dictionary, we can re-write Eq. 1.10 as,

$$\mathbf{x} \approx \sum_{i=1}^{n} \boldsymbol{\alpha}_i \mathbf{d}_i = \boldsymbol{\alpha}\mathbf{D}. \tag{1.11}$$

So, sparse representations in image processing deals with the following two major steps:

(a) **Learning**: Given the training dataset $\mathbf{X} = \{\mathbf{x}^j\}$, $j = 1, 2, ..., m$ is an ensemble of similar patch vectors collected from similar images that learns the dictionary $\mathbf{D}$ and coefficients $\boldsymbol{\alpha}$.

Dictionary learning is carried out by formulating the following unconstrained

optimization problem:

$$\arg \min_{\{\mathbf{d}_i\},\{\boldsymbol{\alpha}^j\}} \sum_{j=1}^{m} \left\| \mathbf{x}^j - \sum_{i=1}^{n} \boldsymbol{\alpha}^j \mathbf{d}_i \right\|^2 + \lambda \sum_{j=1}^{m} \sum_{i=1}^{n} \left| \boldsymbol{\alpha}_i^j \right| \\ \text{subject to } \|\mathbf{d}_i\|_2^2 \leqslant c, \ \forall i = 1, 2, ..., n. \tag{1.12}$$

Here, λ is the regularization parameter. If $\mathbf{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_n]$ represents the matrix consisting all the sparse coefficient vectors, the above optimization problem in matrix form can be written as:

$$\mathbf{D} = \arg \min_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_1 \\ \text{subject to } \sum_i \mathbf{D}_{i,j}^2 \leqslant c, \ \forall i = 1, 2, ..., n, \tag{1.13}$$

where the symbol F in the first term indicates the Frobenius norm where it enforces ℓ_2-norm constraints to the columns of the dictionary $\mathbf{D}$ and the second term finds the ℓ_1-norm $\|\mathbf{A}\|_1$ enforcing the sparsity constraint [115]. Its an optimization problem with two unknown variables and not convex in both $\mathbf{D}$ and $\mathbf{A}$. For solving the optimization problem, it can be split into two reduced subproblems by keeping one fixed and updating the other in an alternative manner [42]. Different techniques are available for solving the dictionary learning and coefficient estimation problem from such two variable regularization problems e.g. coupled dictionary learning via efficient sparse coding [56], separable dictionary learning [51], etc.

(b) **Encoding**: Given the test data $\mathbf{x}$ and dictionary $\mathbf{D}$, computes the sparse coefficients $\boldsymbol{\alpha}$.

Encoding, also known as *sparse coding*, solves an unconstrained quadratic convex optimization problem using a learned overcomplete dictionary $\mathbf{D}$ and vector input data $\mathbf{x}$ as given below.

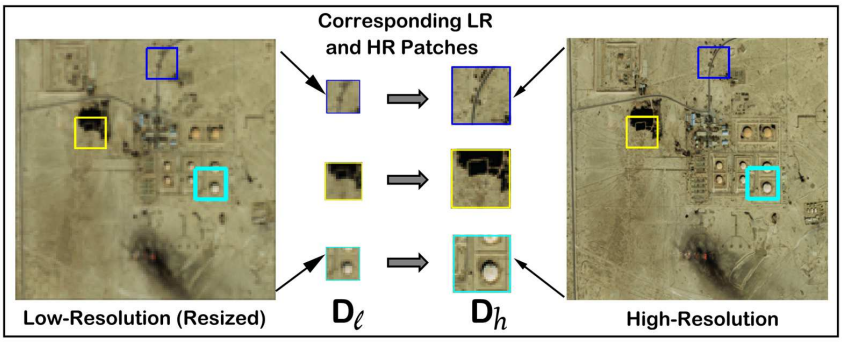
$$\boldsymbol{\alpha} = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \tag{1.14}$$

This is a linear regression problem with the coefficients regularized using the

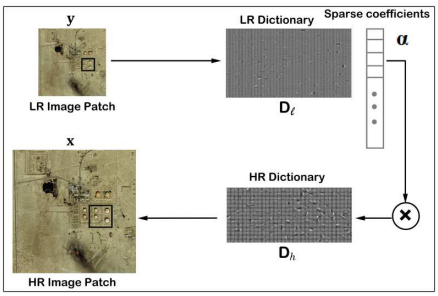
ℓ_1-norm. Many solvers are available for solving such sparse approximation problems e.g. coordinate descent method [47], interior point methods [79], fixed-point-based method [50], iterative shrinkage-thresholding algorithms [8, 9].

1.7 SR model using sparse representation

Given an LR image $\mathbf{Y}$, sparse representation-based SR model tries to recover the HR image $\mathbf{X}$ by defining two constraints, namely, the *sparsity prior* and *global reconstruction* constraints. The method begins by learning two dictionaries, $\mathbf{D}_\ell$ and $\mathbf{D}_h$, from the training images. The LR dictionary $\mathbf{D}_\ell$ is trained by extracting high-frequency feature patches from LR training images, whereas the HR dictionary $\mathbf{D}_h$ is trained by using actual HR patches.



(a) Correspondence between LR and HR patch pairs



(b) Generation of HR patch

Figure 1.10: Steps of patch-wise sparse representation: (a) example of correspondence between LR and HR patch pairs, (b) generation of HR patch using the sparsity model.

The HR image is reconstructed by solving the two regularization problems for-

mulated using the above two constraints and utilizing the learned dictionary pairs. The detailed procedure is as explained below.

(i) ***Sparsity prior-based constraint***:

The sparsity prior assumes that the LR and HR patches extracted from the same location of an LR-HR image pair share a sparse coefficients vector. Fig. 1.10(a) shows the correspondence of LR and HR pairs of an image. First, sparse coefficient vectors are computed for each LR image patch $\mathbf{y}$ by solving a sparse representation problem with LR dictionary $\mathbf{D}_\ell$ is as given below:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ subject to } \mathbf{y} = \mathbf{D}_\ell \boldsymbol{\alpha} \tag{1.15}$$

Then a target HR patch $\mathbf{x}$ can be generated by multiplying the HR dictionary $\mathbf{D}_h$ with above computed $\boldsymbol{\alpha}$ of Eq. 1.15 as follows:

$$\mathbf{x} = \mathbf{D}\hat{\boldsymbol{\alpha}}. \tag{1.16}$$

Fig. 1.10(b) pictorially demonstrates the sparsity based generation of a HR patch using a pair of trained LR-HR dictionaries.

(ii) ***Global reconstruction-based constraint***:

The HR image $\mathbf{X}_0$ obtained by organizing all the HR patches $\mathbf{x}$ generated using the sparsity constraint lacks in patch consistency and it is further regularized using the global image reconstruction constraint i.e. $\mathbf{Y} = SH\mathbf{X}$ which is defined as:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|SH\mathbf{X} - \mathbf{Y}\|_F \text{ subject to } \mathbf{X} = \mathbf{X}_0. \tag{1.17}$$

Here, S and H represents a blurring and downsampling operator, respectively. Finally, by solving the quadratic problem below, an approximation of the target HR image is obtained.

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|SH\mathbf{X} - \mathbf{Y}\|_2^2 + c\,\|\mathbf{X} - \mathbf{X_0}\|_2^2\,, \tag{1.18}$$

where c is the regularization parameter. Here, back-projection of imaging constraint into the patch based HR image helps in inducing the original image's blurring or smoothing properties into the target SR output image and thereby reducing the patch inconsistencies.

1.8 Motivation of the present work

Sparse representation and compressive sensing provide a reliable statistical framework for analyzing high-dimensional data and methods for revealing data structures, resulting in a substantial repertoire of efficient algorithms. Greedy methods and convex optimization can correctly and efficiently compute a sufficiently sparse linear representation [7]. New formulations based on sparse representation allow to explore and reveal more meaningful structures/features for various data and propose efficient optimization strategies for such data. We know that MS images suffer from low spatial resolution which has been a bottleneck for many HR applications. Generation of HR MS image(s) from a single LR MS image through SR processing will be highly beneficial in remote sensing image analysis and applications.

From, literature it is studied that, sparse representation-based SISR of natural images can reproduce the missing fine elements in LR images by learning appropriate dictionary. However, due to the limited spatial resolution of remote sensing MS images, building a good quality overcomplete dictionary for SR applications is problematic. Furthermore, existing SISR algorithms are not easily applicable to actual MS images. Sparse representation methods have the flexibility to develop new strategy using recent advancements in dictionary learning or sparse representation phase. Recently, examples of sparse representation applied for multidimensional signal or image processing can be found in the literature. This has motivated the researcher to apply recently developed sparse representation methodologies and convex optimization tools to the MS image SR problem.

Image processing algorithms are suitable for parallelization as they have high

level data redundancy. A single instruction is executed iteratively over a million of pixels or large number of patches in a loop. With recent development of high performance computing facilitates such computationally expensive algorithms can be implemented using parallel processing. Since, real-time remote sensing operations need to process images of large dimension, this motivates to implement the proposed MS image SR algorithm using parallel processing.

1.9 Contributions from the Thesis

The research works carried out in the thesis result in the following contributions:

i) Developed a multicore SISR method using trained sparse overcomplete dictionaries. Also proposed new dictionary training strategies using MS remote sensing images and developed a SR reconstruction algorithm (named it as 'MSISR') using the concept of patch-wise sparse representation. Furthermore, for MS datasets that include HR panchromatic images as well as LR MS bands, a new spatial-spectral SR influenced by the pansharpening technique is proposed to reconstruct HR MS images from only the LR MS image. The proposed method has been found advantageous over pansharpening-based fusion techniques.

ii) Developed a new MS image SR method using the morphological component analysis (MCA)-based features and sparse representations. The proposed work also uses principal component analysis (PCA) to demonstrate a new dictionary learning methodology from a set of MS images. Moreover, fast implementation of both the dictionary training as well as SR reconstruction is carried out by utilizing the Open Multi-Processing (OpenMP) parallelization and achieved a reasonable speed-up. Reconstructed images have shown better remote sensing analysis, when subjected to spectral signature evaluation and end-members identification from the super-resolved images.

iii) Proposed a new joint sparse representation based MS image SR algorithm

(named it as 'JAMiSR') by combining the patch and patch-groups in a common framework. Also, proposed a self-adaptive dictionary learning approach for MS images. Extensive comparisons of results are shown with the state-of-the-art sparse representation- and some deep learning-based SR algorithms. To make it suitable for near real-time remote sensing applications, the proposed algorithm has been implemented on general-purpose graphics processing units (GPGPU) and results are demonstrated. Furthermore, land-cover classification is performed on the super-resolved images images to show the potential of the proposed method for real remote sensing applications.

List of publications:

A. *Journals:*

(i) Bhabesh Deka, Helal Uddin Mullah, Trishna Barman and Sumit Datta, "Joint Sparse Representation-based Single Image Super-resolution for Remote Sensing Applications", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2352-2365, February 2023.

(ii) Helal Uddin Mullah and Bhabesh Deka, "Fast Multispectral Image Super-resolution via Sparse Representation", *IET Image Processing*, vol. 14, no. 12, pp. 2833–2844, May 2020.

B. *Book Chapters:*

(i) Helal Uddin Mullah, Bhabesh Deka, Trishna Barman, and A.V.V. Prasad, "Sparsity Regularization Based Spatial-Spectral Super-Resolution of Multispectral Imagery", In: Deka B., Maji P., Mitra S., Bhattacharyya D., Bora P., Pal S. (eds.) Pattern Recognition and Machine Intelligence (PReMI 2019). *Lecture Notes in Computer Science, Springer*, vol. 11941, pp. 523–531, November 2019.

(ii) Helal Uddin Mullah and Bhabesh Deka, "A Fast Satellite Image Super-Resolution Technique Using Multicore Processing", In: Abraham et al.

(eds.) Hybrid Intelligent Systems (HIS 2017). *Advances in Intelligent Systems and Computing, Springer*, vol. 734, pp. 51–60, March 2018.

C. *Conferences:*

(i) Helal Uddin Mullah and Bhabesh Deka, "Parallel Multispectral Image Super-resolution based on Sparse Representations", *IEEE International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*, February 2019.

1.10 Scope of the work

MS images are widely used for different remote sensing applications; reconstruction of high quality images is of paramount interest. Keeping in mind about the limitations of HR MS image data availability; challenge is to apply the traditional sparse representation-based SR methods for MS datasets. Learning an effective dictionary that can enhance the performance of sparse representations and effectively restore HR MS image is very important. It is noticed that learning a global dictionary from similar HR images collected from some external databases for many LR imaging satellites is a challenge. This leads to search for novel techniques and strategies of MS image dictionary learning.

Moreover, we observe that an SR method capable of producing the target HR image directly from the input LR image will be more practical because it is not reliant on an external HR image. Such an adaptive SR algorithm capable of producing state-of-the-art results (both visually and in remote sensing aspects) will be suitable for practical remote sensing applications as long as the computation time is reasonably close to that of near real-time operations. Therefore, parallel processing-based implementations of novel MS image SR algorithm on a high computing facility may find a place in the real remote sensing applications. For example, a ground station which collects LR MS images can consider such a system for real time visualization of corresponding HR images, or an aircraft integrated with LR camera can embed an

SR hardware module for real-time use; research organizations or companies dealing with LR satellite image data can consider such a parallel SR imaging system to avail the benefits of high spatial and spectral reconstructed MS images. Overall, it can be stated that the proposed works in this thesis have ample scopes in the practical remote sensing applications.

1.11 Thesis outline

The thesis is organized into six chapters. In the following, a brief introduction to each chapter is given:

Chapter 1:

A brief introduction to image SR utilizing sparse representation is provided in this chapter. The principles of image SR, SR in remote sensing, its benefits and drawbacks, the concept of sparse representation, and how sparse representation solves the inverse problem of image SR are all covered in this chapter.

Chapter 2:

This chapter presents a comprehensive review of the traditional super-resolution methods, sparse representation based SR methods and some widely used pansharpening methods. We discuss the different SR techniques, their performances and limitations in obtaining the HR image. A brief background on parallel processing in image processing and review of a few related parallel processing-based SR works are also presented in this chapter. The chapter concludes with a brief summary and highlighting few research issues on this topic.

Chapter 3:

In this chapter, we present three pansharpening-based fast MS image SR using sparse representations. First method demonstrates a fast satellite image SR technique using multicore parallel processing for RGB satellite images, where SR algorithm is applied on the luminance channel after YCbCr colour transformation. Coupled HR/LR dictionaries are trained from external HR image datasets using the joint sparse coding technique. Next, we develop a new dictionary learning scheme by creating an

HR training dataset from patches of multiple PAN images and an LR dataset from concatenated patches of corresponding LR MS images. We also propose MSISR, a sparse representation-based multicore parallel MS image SR reconstruction algorithm. In the third method, a pansharpening-based SR algorithm is developed using a single pair of PAN and LR MS images, where a coupled dictionary is learned from the PAN image itself. This method helps in restoring better spatial as well as spectral information compared to some of the existing pansharpening-based methods.

Chapter 4:

This chapter describes a new SR method for reconstruction of MS images collected by the Indian remote sensing satellite Resourcesat-2. We identify the essence of morphological feature extraction from MS image for effective sparse representation and also proposed a PCA led dictionary training technique for images collected by Linear Imaging and Self Scanning Sensors (LISS), namely, LISS-III and LISS-IV, respectively. Multicore parallel algorithms are designed using OpenMP; achieves significant speed-up both in dictionary training and SR image reconstruction.

Chapter 5:

In this chapter, we propose an adaptive SR algorithm exploiting the benefits of joint sparse representation for MS images super-resolution of Indian satellite as well as publicly available remote sensing images. The method named as 'JAMiSR' utilizes both patch- and patch-group-based sparse representation and able to produce a better image compared to the state-of-the-art sparse- and deep learning-based methods. Furthermore, we accelerate the proposed algorithm using NVIDIA GPGPU parallel processing hardware so that it can be used in real-time applications.

Chapter 6:

This chapter brings the thesis to a close by summarizing the works done and outlining some potential future research in the same field.