Chapter 2

Review of Literature

2.1 Introduction

Even though the imaging sensors are being continuously upgraded, there are some physical constraints still, which motivates the computer vision research community to address them with the help of efficient signal or image processing algorithms. Super-resolution (SR) is one such image processing technique that deals with low-resolution (LR) images. However, only a few SR works that focus on improving the resolution of multispectral (MS) remote sensing images are available in the literature. In this chapter, we provide a brief history and taxonomy of available SR methods, highlighting their benefits and drawbacks in the context of MS image SR.

2.2 Taxonomy of image SR methods

In the literature, many authors have presented topy surveys on image SR methods pertaining to different applications [22, 41, 73, 76, 103]. So, their classification is not uniform across all the available literature. Here, we present a generalized taxonomy of major SR methods, selecting only a few significant works for discussion from each category.

Fig. 2.1 gives an overview of different SR approaches. It is classified into two types: *frequency domain* and *spatial domain methods*.

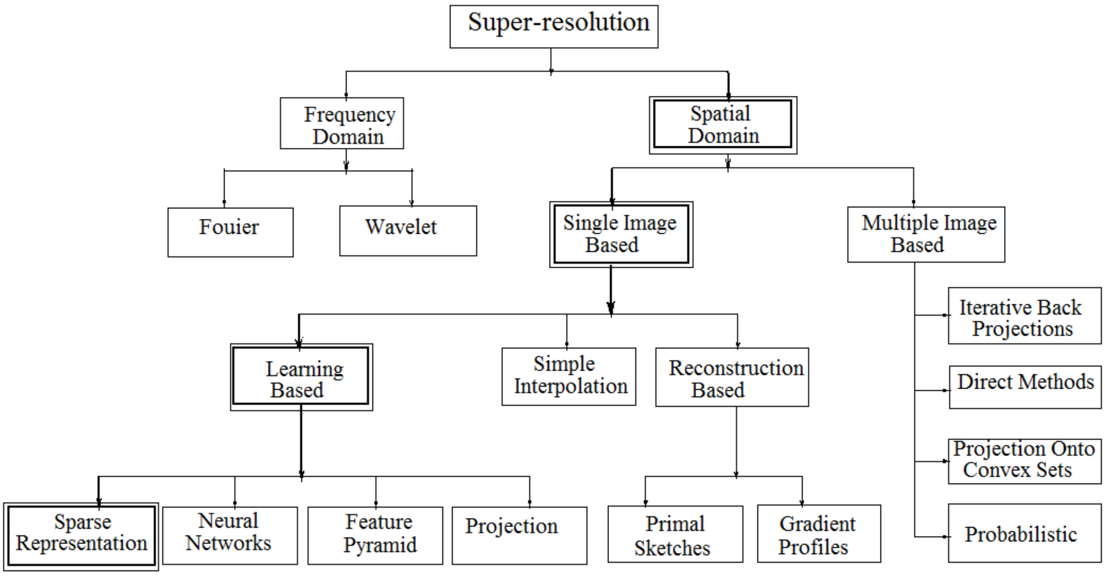

Figure 2.1: Taxonomy of some major SR methods in the literature

2.2.1 Frequency domain SR methods

Tsai and Huang [104] developed early methods based on the shifting property of the **discrete Fourier transform** (DFT) to generate sub-pixel shifted versions of the given noise-free LR image. The shifting property relates spatial domain translation to frequency domain phase shifting and is denoted as follows:

$$F_s\left(\mathbf{u},\mathbf{v}\right) = e^{j2\pi(\mathbf{u}\triangle\mathbf{x}+\mathbf{v}\triangle\mathbf{y})}F\left(\mathbf{u},\mathbf{v}\right), \tag{2.1}$$

where $\triangle\,\mathbf{x}$ and $\triangle\,\mathbf{y}$ signify the subpixel shifts in the image's x- and y-directions, respectively. The inverse DFT of $F_s\left(\mathbf{u},\mathbf{v}\right)$ will give the corresponding sub-pixel shifted LR image $f\left(\mathbf{x}+\Delta\mathbf{x},\mathbf{y}+\Delta\mathbf{y}\right)$ from the available image $f\left(\mathbf{x},\mathbf{y}\right)$ [76]. Different sub-pixel motions provide complementing information across LR frames, allowing for SR reconstruction [107]. The **wavelet transform**, which decomposes the input image into structurally correlated sub-images to retrieve high-frequency information, is another approach in this category reported in [16, 27, 32]. These methods exhibit theoretical simplicity and are straightforward for implementation. They are computationally efficient too. However, they fail to give good quality results in real-world applications due to poor translational models, degraded convolution filters, and lack of prior knowledge.

2.2.2 Spatial domain SR methods

These methods do not transform the given image into frequency domain, instead estimate the the high-resolution (HR) image using techniques that operate in the spatial domain itself. The advantages of spatial domain methods are that they support unconstrained motion between frames and have scope for easy incorporation of prior knowledge. These methods are divided into two types based on the number of LR images: *multiple image SR* and *single image SR*.

2.2.2.1 Multiple image SR methods

The **iterative back projection** (IBP) algorithm was proposed by Irani *et al.* [54] for multiple image SR. The HR image is determined by back projecting the differences between the simulated and observed LR images, as illustrated in Fig. 2.2. The method starts with a rough estimate of the HR image, then iteratively repeats the projection for each observed LR image, while minimizing the error term $E(\mathbf{x}_H)$-,

$$E(\mathbf{x}_H) = \frac{1}{2}\left\|\mathbf{x}_L - A(\mathbf{x}_H)\right\|_2^2, \tag{2.2}$$

where $\mathbf{x}_L$ and $\mathbf{x}_H$ represents the observed LR and the estimated HR images, respectively. This method is simple, but fails to deal with noise associated with LR images and its convergence is not guaranteed. Moreover, *a prior* constraint is not available in the IBP.

A **direct method** proposed by Chirang *et al.* [25] produces HR image based on warping and registration of already scaled-up LR images. It is faster than IBP. **Projection on Convex Sets** (POCS) is another multiple image SR method by Stark *et al.* [95] that incorporates prior knowledge of a closed convex set for each LR image and iteratively estimates the HR image. This method suffers from slow convergence and high computational costs. For multiple image SR, **probabilistic-based methods** such as Maximum Likelihood (ML) and Maximum a Posteriori (MAP) estimate the HR images by minimizing ML (or MAP) cost functions, while

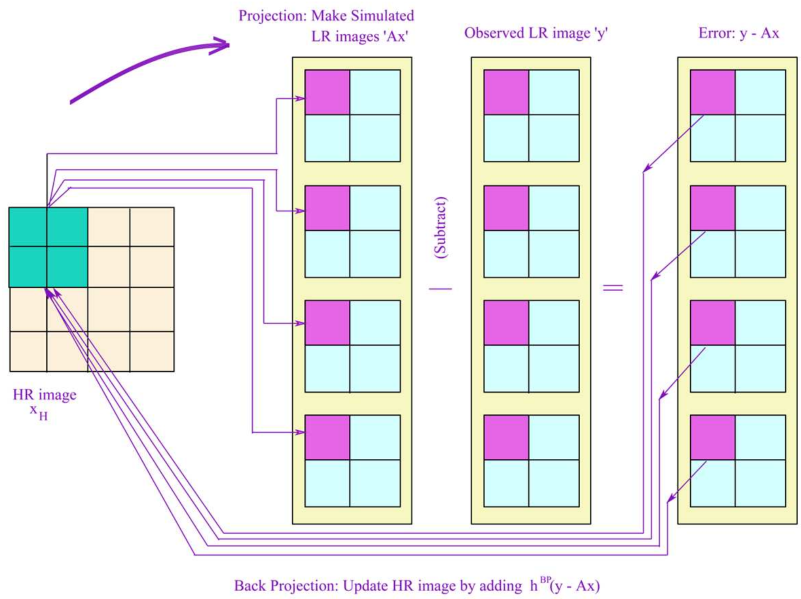

Figure 2.2: Overview of IBP super-resolution method

restricting the solution within specified sets. For insufficient number of inputs they provide good results by involving prior knowledge [21, 77]. In real-world applications, where various objects within the same image can have varied and complex motions; multiple image-based SR methods depend largely on the accuracy of motion estimation between LR images, which is highly unreliable.

2.2.2.2 Single-image SR methods

There are **simple interpolation**-based techniques, like, linear, bicubic, etc. for single image SR (SISR). However, they have limitations in terms of blurring and fail to enrich the high-frequency information in the HR image as no prior knowledge about the target image is available during reconstruction [45]. **Reconstruction-based methods** do not employ a training set and instead rely on statistical image priors to increase reconstruction quality. Primal Sketches-based methods use primal sketch as *a priori* information pertaining to the primitive parts (e.g. edges, corners, ridges, etc.) of the LR image in a patch-wise manner to produce an HR image [100]. **Gradient profile** priors, like general Gaussian distribution (GGD) [98, 99], gradient profile sharpness [114], etc. are also used to utilize the shape of gradient profiles in LR and HR images as reconstruction constraints.

To solve the SR problem, **learning-based methods** have recently been introduced, which use a trained dictionary to estimate the target HR image by learning co-occurrence between LR and HR patches from an external LR-HR database. By collecting the most likely high-frequency information from the training images, the high-frequency information present on the LR image is enhanced based on local features. Fig. 2.3 depicts a standard structure for learning-based SR methods..

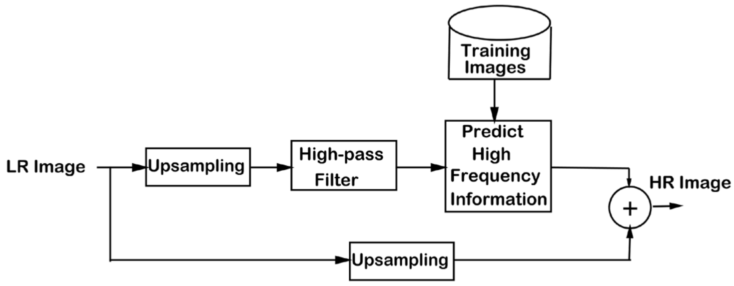

Figure 2.3: A framework for the SR imaging methods based on learning

Some algorithms use **projection based** learning of the a priori term as done by Capel *et al.* [15]. Miura *et al.* [68] presented a principal component analysis (PCA)-guided learning-based SR method. **Feature pyramids** are learned from training images e.g. Gaussian pyramids containing low-frequency information, and Laplacian pyramids containing band-pass information of face images are applied as prior information during the SR reconstruction [97]. In SR, various types of **neural networks** are also used to learn an end-to-end mapping between LR and HR images from large datasets in order to estimate the parameters of the target HR image. Deep learning techniques, like convolutional neural networks (CNN)-based algorithms are popular in this category [61]. According to recently developed sparse representation theory (SRT) an image patch can be represented as a sparse linear combination of elements from an overcomplete dictionary [37]. An LR patch can be effectively represented using a pair of jointly-learned LR-HR dictionaries to form a **sparse representation** problem; the solution of it is being used for the generation of corresponding HR patch [115]. With recent developments in efficient convex optimization methods and high-end computing facilities, research on sparse representation-based SR methods are largely carried out for various applications.

2.2.3 Pansharpening methods

The spatial information in the HR grayscale PAN image, and the spectral information in the LR MS bands may be combined to produce high spatial/spectral MS images. These techniques are known as pansharpening because they equalize the resolution of MS and PAN images. In [4], a comprehensive review of works on the fusion of MS and PAN bands is carried out. The intensity-hue-saturation (IHS) [106], principal component analysis (PCA) [90], and Brovery transform-based methods [128] are three prominent pansharpening methods. The MS image is first transformed into RGB and then resized to the size of the PAN image. The resized RGB image is then transformed into the YCbCr format and the Y-channel (luminance) is replaced by the PAN image. Lastly, inverse transformation is done to get the RGB MS image back from the YCbCr image. One notable drawback of the pansharpening techniques is that the resulting images seem to have severe spectral irregularities. The reason for this is that the pixels in the luminance channel do not have the same statistical distributions as those of the grayscale PAN image [91].

In this thesis, we will use the SISR technique to reconstruct HR-MS images with enhance spatial information. The pansharpening and SR techniques can be distinguished because the former attempts to produce a MS image with the spatial resolution of a PAN image, whereas the latter attempts to simultaneously enhance the spatial and spectral resolution of a given LR MS image.

2.3 Sparse representation-based SR methods

Sparse representation, often known as sparse coding, is a powerful image modeling technique for solving the inverse problem of image restoration [63, 64]. A basic sparsity-based SISR model consists of the dictionary training and image reconstruction phases [2]. The given training data, i.e. a group of LR and HR patch pairs, is used to jointly learn a pair of LR and HR dictionaries. In the reconstruction step, the trained dictionaries are used to encode each overlapping patch of theLR inputimage, and a sparse coefficient vector is obtained by solving a convex optimization problem. By multiplying these coefficient vectors with the HR dictionary, the relevant HR patches are generated. The final HR image is produced by registering and aggregating all the HR patches into a grid of target size. Many SR methods that use the sparsity prior have been proposed as a result of recent breakthroughs in sparse coding and dictionary learning techniques. In the subsections that follow, some sparse representation-based methods are briefly addressed.

2.3.1 Patch sparse representation-based dictionary learning and SR image reconstruction

Yang *et al.* [115] first proposed the sparse representation-based SR method called the sparse coding super-resolution (ScSR) that proves the application of sparse representations for the reconstruction of HR natural images. It is essentially based on learning a coupled overcomplete dictionary trained over a large dataset of LR and HR image patch pairs for the sparse representation of the test LR patches. The sparse coding problem is solved using the 'feature-sign search' [56] -based ℓ_1-norm minimization technique. While, in the reconstruction, the sparse coefficients are multiplied by the trained HR dictionary to reconstruct the corresponding HR patches. The use of high-pass filter-based gradient feature extraction improves the reconstructed image details, while a back-projection step increases the consistency among the HR patches by reducing the aliasing effect. Although ScSR has achieved reasonable success in improving the images both visually and quantitatively, yet they lack in minimizing the serrated edges and ringing artifacts. Moreover, the use of global dictionaries to represent any input image may not be effective as they are not adaptive and chances of misinterpretation is also high.

Authors in different works provided approaches to improve the results of ScSR either by learning an effective dictionary on high-frequency feature vectors or adding new regularization constraints with novel *a priori* information in the reconstruction problem. Dong *et al.* [35] propose an image deblurring and super-resolution technique using the sparse representation model. First, k-means clustering is done to classify the training image patches into several classes and then learns different compact overcomplete dictionaries from them. During reconstruction of a patch, it adaptively selects the most relevant dictionary to characterize the local sparsity. Experiments are conducted on natural images to get reasonable improvements. However, the algorithmic complexity of the method is high due to the presence of three penalty terms, which is not preferred for fast SR. A similar approach for multi-dictionary learning-based compressive sensing (CS) SR of synthetic aperture radar (SAR) images is proposed by He *et al.* [52]. CS-based SR methods are also explored along with self-similarity regularization for improved performances [30, 75].

Zhu and Bamler [133] propose a sparse representation-based MS image fusion method named as 'SparseFI' applying the HR PAN image information. Here, coupled dictionaries are directly generated from the PAN image and use to restore the HR MS bands. It is to be noted that the available MS images are not used for dictionary learning here, which causes spectral loss in the reconstructed images. To overcome this, Li *et al.* [59] propose a remote sensing image SR through coupled dictionaries learned using the PAN and LR MS images adaptively. Then constructs the dictionary for an unknown HR MS image by using these dictionaries. In another work, Guo *et al.* [49] propose an iterative method to first combine the PAN and the LR MS image to get an HR MS image and then learn a HR dictionary to be used in the next iteration. This method reduces SparseFI's limitations by training dictionary from intermediate MS image instead of PAN image. The proposed method overcomes the spectral distortion of fusion methods by incorporating the MS image into the dictionary training stage.

There are also sparse representation-based SR works that have exploited new feature extraction schemes for improved results. Chavez-Roman and Ponomaryov [20] have done wavelet domain interpolation with edge extraction for the estimation of high-frequency sub-bands. Similarly, Yang *et al.* [118] propose remote sensing image SR methods based on learning of primitive- and residual-sparse dictionaries to recover primary and residual high-frequency information. Multiple features e.g.

gradients, histogram of oriented gradients (HoG), Gabor features, etc., are extracted to describe the image structures.

ScSR and SR methods based on ScSR, process individual patches ignoring the consistency of pixels in overlapped patches, which is a strong constraint for image SR. Alvarez-Ramos *et al.* [2] propose a satellite image SR method based on sparse representation of overlapping patches, where PCA is applied to reduce the dimension of feature patches and K-SVD-based dictionaries are trained. Gu *et al.* [48] develop the convolutional sparse coding based image super-resolution method named as 'CSC-SR', which works on the whole image, does not need to divide the image into overlapping patches. This method can exploit the global correlation for the robust reconstruction of local structures.

Moustafa *et al.* [70] propose a hyperspectral image SR method using the compressive sampling matching pursuit (CoSaMP) algorithm. PCA-based significant bands are selected for dictionary training and sparse representation, less significant bands are upscaled through interpolation. They have shown real-time speedup using CUDA programming model and the cuBLAS library. In another work [71], they have shown MS image SR via self example learning and sparse representation. Here, they demonstrate an adaptive dictionary learning from the LR image and obtained speedup of 10-20$\times$ for image sizes up to 512$\times$512, and 20-40$\times$ up to 2048$\times$2048 images by MEX-CUDA based implementation.

In several works [3, 82], wavelet preprocessing is utilized to create four sub-bands from an input image, namely low frequency approximation, horizontal, vertical, and diagonal detail bands, and focuses on the training of features obtained from these four subbands. Rapid and accurate image super-resolution (RAISR) is based on image sharpening and enhancement by amplifying the underlying details using a set of filters that may be learned [85].

The non-local similarity property of image patches is effectively utilized as prior information for improved SR reconstruction by the sparse representation methods [18]. Chang *et al.* [17] combine the complementary collaborative sparse representation-based regularization (CRR), and the non-local low-rank regularization (NLR) such that both external and internal HR information are well preserved. The method is named as the ‘collaborative representation and non-local self-similarity (CRNS)’ and results are demonstrated for medical image SR. Chen *et al.* [23] introduce an optimization method for non-convex and non-separable regularization problems for simultaneous sparse and low-rank matrix reconstruction. Experiments are shown with synthetic as well as real hyperspectral images. Similarly, Rencker *et al.* [83] propose dictionary learning and compressive sensing recovery for non-linear measurements e.g. clipped or quantized measurements using the non-convex optimization based approach. Shao *et al.* [92] propose a coupled sparse auto-encoder (CSAE) as an alternative to the joint dictionary training methods, where the HR coefficients predicted by their method exhibit larger correlation with the true values.

2.3.2 Group sparse representation-based SR image reconstruction

Traditional patch-based SR imaging faces major difficulties as it must solve computationally expensive large-scale optimization problems for dictionary training and sparse reconstruction. Also, overlapping patches are processed independently in both stages. As a result, the spatial correlation among patches is overlooked, resulting in an imperfect approximation or an unstable sparse representation. However, it is observed that similar patches are usually present at multiple spatial locations within the image irrespective of any scale [45]. Although different SR works combining the non-local self-similarity (NLSS) and patch sparsity [17, 75] properties are available, only a few works are reported till date exploiting the group-sparse representation (GSR) for remote sensing image SR.

To address the above issues, Zhang *et al.* [126] propose a GSR-based image restoration method that can enforce both inherent local sparsity and non-local self-similarity at the same time. In addition, instead of learning coupled overcomplete dictionaries, low-complexity self-adaptive group dictionaries are learned for each group. Examples of GSR in image restoration, denoising, despeckling, and compressive sensing recovery can be found in the literature [58, 62, 127]. It is observed that although GSR provides a traceable solution in the same scale, but for SR, the reconstructed images with higher upscaling are effected by oversmoothing.

Xu and Gao [113] propose a SR method by directly adopting the GSR technique to upscale an image by 2. Normally, non-local similar patches are selected based on the Euclidean distance among the patches. Here, the authors considered the Gaussian kernel distance instead of Euclidean distance to better represent the geometrical structures of the image. Liu *et al.* [62] utilize the concept of GSR to explore the underlying patterns of SAR images.

Recently, joint patch-group sparse representation (JPG-SR) techniques combining the patch sparse representation (PSR) and group sparse representation (GSR) are greatly explored for improved SR performances in natural image restoration problems, like, inpainting, deblocking, etc. [124]. Owing to the fact that PSR generates undesirable visual artifacts, while GSR model tends to show oversmoothing effects, the JSR tries to integrate the local sparsity and non-local self-similarity of images [123]. The JPG-SR model performs an alternating direction method of multipliers (ADMM)-based optimization of the joint regularizations involving PSR and GSR constraints. Mikaeli *et al.* [66] propose a single-image SR for natural images via patch- and group-based local smoothness modeling (SR-PGLSM). They adopt the isotropic total-variation technique for modeling of patch-based local smoothness. A complementary regularization term based on non-local means is considered to develop (SR-PGLSM-NLM) algorithm and finally solves it using the split Bergman iterative technique [46]. In another work, Gao *et al.* [44] develop the joint sparse and low-rank learning (J-SLoL) for enhancing the spectral information of MS images using partially overlapped hyperspectral images. They also apply the ADMM method for optimization of the proposed J-SLoL algorithm. They have evaluated their algorithm based on sparse reconstruction, classification, and unmixing of the generated images.

2.4 Parallel computing for image super-resolution

Super-resolution involves processing big data, requiring significant computational and memory resources for obtaining the target image. A 1280×720 (HD quality) colour image contains 9,21,600 numbers of pixels in each band and occupies 27, 64,800×8 ≈ 22 MB of total memory space. Similarly, data volume is potentially large for the satellite images as well, and some SR techniques e.g. multiple image registration-based SR will usually increase the required computations. Moreover, most of the dictionary learning-based SR methods require a large dataset of training images to train the LR and HR dictionaries, and it solves several regularization problems, which are computationally intensive, to tackle the ill-posed inverse problem of SR image reconstruction.

One of the most recent innovations in computational machines is the development of multicore processors, which consists of two or more independent cores in a single package. Many processors nowadays incorporate multicore architectures to meet the rising demand for greater performance. Most CPU manufacturers currently prioritize improving on-chip multi-threading functionality by increasing the number of cores over raising the processor actual clock speed. By exploiting these hardware advances, we may benefit dramatically by modifying an existing single-threaded code into a multi-threaded one to run on multiple cores. Since image processing algorithms shows both data and instruction level parallelisms (i.e. DLP and ILP), we may modify an existing sequential code into a nested-loop-based parallel code, which can be implemented in a many core or multi-core processor [26]. The speed-up of a particular implementation is determined as the ratio of sequential execution time to parallel execution time of the same algorithm.

2.4.1 OpenMP-based parallel implementation

Open multi-processing (OpenMP) is an application program interface (API), which can be utilized to perform parallel processing through shared memory-based multi-threading operations. OpenMP supports many functionalities required for parallel programming. It is a set of pre-processor directives, runtime library routines, and environment variables that the programmer can use to tell the compiler how to apply multi-threading on a block of code. Since it is platform-independent, properly written OpenMP code for one platform can be readily recompiled and executed on another. Moreover, OpenCV software supports large numbers of in-built functionalities for image processing operations and more importantly it includes the OpenMP-based parallel framework.

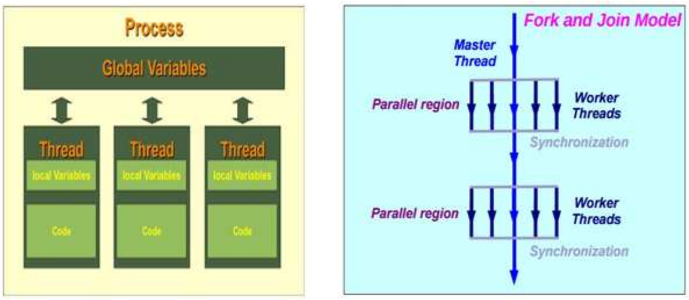

Figure 2.4: FORK-JOIN model for parallelization

The OpenMP specifications are used to instruct the parallel sections of the code to be executed concurrently on different cores of the same processor using the shared memory multi-threading concept. It works according to the FORK and JOIN model, where a sequential program starts in a master thread and then it is divided into some worker threads in the loop sections (FORK), which are finally combined after completion of the loop (JOIN) to get the output. The master thread continues until the whole execution of the code is completed.

Different threads have a common access to the global variables, while the local variables can only be handled by the host thread. In OpenMP, transfer of data is always transparent to the programmer. The runtime environment specifications of OpenMP provides the required number of threads on which the loop should be executed concurrently. Since the master node continues till the end so we should assign some task in the main program for master node also, besides assigning to the different threads.

Examples of works applying multicore computing with OpenMP into image processing is found in the work presented by Greg *et al.* [93], where they show speedup of about 3 to 4 times for different image processing algorithms.

2.4.2 GPGPU-based parallel implementation

With the recent advances of many-core general-purpose graphics processing units (GPGPUs) by NVIDIA Graphics and compute unified device architecture (CUDA) programming interface, there is a high demand for performing machine learning-based remote sensing image analysis [12, 70, 80]. Tan *et al.* [101] apply LASSO for sparse approximation and perform CUDA-based parallel implementation to achieve 30-35 speed-up. A similar work is also presented by Attarde *et al.* [5] for natural image SR. Moustafa *et al.* [71] demonstrate a GPGPU implementation of learning-based multispectral image SR algorithm with 20 to 40$\times$ speed up for different image sizes. It is observed that most of these works focus on the acceleration of the SR reconstruction, while utilizing a sequentially implemented pre-trained dictionary. Additionally, the available parallel works are mostly based on the PSR approach due to which their reconstructed outputs are not as competitive as other state-of-the-art methods.

2.5 Remote sensing datasets

Most SISR algorithms are based on upscaling of grayscale images. For RGB remote sensing images, typically, most SR works use colour transformation from RGB to YCbCr space to apply SISR on the Y-channel [24]. Ruben *et al.* demonstrates single-frame SR using RGB remote sensing images [41]. Another category of MS image fusion works uses the HR PAN and false color LR MS image data to generate a HR MS false RGB image as discussed in [105], However, the raw MS images are provided in the form of multiple band images (3-10) of varying wavelengths. There are different remote sensing satellites, which provide land-cover MS images of different spatial resolution. For example, the University of Marylands Global Land Cover Facility (GLCF)[1] offers thousands of Landsat scenes and derived data products. The United States Geological Survey (USGS)[2] contains very well-maintained collection of Landsat and Sentinel-2 data. In this thesis, we mostly utilize MS images collected from National Remote Sensing Centre (NRSC)[3], Hyderabad and publicly available datasets from GLCF and the Bhuvan portal[4] (maintained by ISRO). Initially, few experiments are carried out using some freely available sample images collected by ISRO's Cartosat-2 series satellite[5]. The NRSC dataset consists of images captured by two Linear Imaging Self-Scanning (LISS) sensors, LISS-III and LISS-IV of ISRO's ResourceSat-2 satellite. We have collected and used publicly available Quickbird MS images from the GLCF portal for some experiments. Additionally, we have used few benchmark remote sensing datasets, like PatternNet[6], aerial image dataset (AID)[7], UC Merced (UCMD)[8] land use dataset, and the CAVE MS image dataset[9].

2.6 Summary and research issues

We know that the key to single image SR (SISR) is to choose the right prior information and then how we utilize it to compensate for the missing spatial information due to the imaging process. For remote sensing images, multiple-image SR methods are generally not preferred because of their limitations for acquisition of multiple images of the same scene using satellite. Learning-based SISR methods, mainly the sparse representation approach, are widely used for better results. However, there are enough scopes for improvement in these methods by duely addressing the issues encountered either in dictionary learning or sparse representation, or database preparation. From the above literature study, we may summarize the following

[1]http://glcf.umiacs.umd.edu
[2]https://earthexplorer.usgs.gov
[3]https://uops.nrsc.gov.in/
[4]http://bhuvan.nrsc.gov.in/data/download/index.php
[5]http://www.isro.gov.in/pslv-c38- cartosat-2-series-satellite/images-cartosat-2-series-satellite
[6]https://sites.google.com/view/zhouwx/dataset
[7]https://captain-whu.github.io/AID/
[8]http://weegee.vision.ucmerced.edu/datasets/landuse.html
[9]https://www.cs.columbia.edu/CAVE/databases/multispectral/

research issues on remote sensing image SR:

2.6.1 In dictionary learning

i. Dictionary training plays the crucial role for sparse representation-based image SR. Dictionary trained over the high-frequency features of training images can improve the accuracy of reconstruction.

ii. Existing methods mostly learn the dictionary offline due to which it is observed that the reconstruction results are good only for few specific images (similar to the training images). But, we know that for SR using sparse representations, results will be better if the learned bases or dictionary atoms are more relevant to the test data [134]. Therefore, instead of learning a pair of overcomplete dictionaries from external training datasets, we may train an adaptively learned overcomplete dictionary from the test image itself.

iii. Size of the dictionary determines the speed of the SR algorithm and quality of the reconstructed images. There should be a tradeoff among these attributes, when learning an effective dictionary. Other aspects, like- dictionary initialization, feature extraction, optimization approaches for sparse representation, assumptions on the image degradation model, like amount of blurring and noise, etc., are also important for obtaining a robust and effective dictionary.

2.6.2 In SR reconstruction

i. If HR image database is available for dictionary learning, good quality SR reconstruction is possible through extraction of high-frequency features from the underlying image patches as it will represent the image textures rather than the absolute intensities.

ii. Panchromatic images can be utilized as prior information in the SR reconstruction of MS images as both the PAN and HR version of the observed LR MS image represents the same spatial information.

iii. Extra information from identical structures known to occur in remote sensing images can be incorporated into the dictionary learning, which is utilized to reconstruct the target HR image.

iv. Within a unified framework, the group-based sparse representation can enforce both intrinsic local sparsity and non-local self-similarity of images. However, selection of groups which can capture the non-linear nonlocal structure information properly is a matter of concern.

v. It is found that assuming the patch and group of patches for local and non-local prior modeling, joint sparse representation formulation can improve the SR reconstruction results. But, a proper design of the joint SR algorithm and efficient solving method is required for functioning of such methods.

vi. In SR methods, regularization processes, involved for solving the sparse approximation problems, make them slow. Also, the dictionary training is a very time consuming process.

vii. Satellite image data are naturally of large sizes and real-time SR reconstruction for remote sensing applications is quite challenging.

viii. Parallel processing hardware-based techniques for SISR give good accelerations required for near real-time implementations in remote sensing applications.