9 Major Techniques and Algorithms for Multisensor Data Fusion

9.1 INTRODUCTION

With the rapid advancement of space technologies and remote sensing tools/platforms such as the Earth Observing System in the National Aeronautics and Space Administration (NASA), a variety of remotely sensed data have become available in the Earth Observing System Data Information System (EOSDIS) from diversified operational satellites since the 1980s. Nevertheless, there exists a well-known trade-off between spatial, temporal, and spectral resolution of remotely sensed data in various remote sensing systems. In fact, it is always difficult to obtain one image measured by one individual instrument that has both high spatial resolution and high spectral resolution as well as high temporal resolution. For example, remotely sensed imagery from a satellite-based sensor like Landsat Thematic Mapper (TM) has a spatial resolution of 30-m with a temporal resolution of 16-days, while a Moderate Resolution Imaging Spectroradiometer (MODIS) image is 250-m (spatial resolution) on a daily (temporal resolution) basis. Thus, the remotely sensed imagery from TM has a higher spatial resolution than MODIS, but a lower temporal resolution (16-days versus 1-day). This is mainly due to competing constraints between the instantaneous field-of-view and signal-to-noise ratio limited by the detector instruments and satellite orbits.

In real-world applications, simply making use of one type of remotely sensed data may not fulfill all the requirements necessary to better monitor the changing Earth environment. For example, some satellite sensors providing multispectral images identify features spectrally rather than spatially, while other satellite sensors providing high spatial panchromatic images characterize features spatially instead of spectrally (Wang et al., 2005; Chen et al., 2008). These constraints significantly limit the broad exploration of valuable data sources in various applications, especially for those emergency response events that demand high spatial and high spectral (or high temporal) resolution simultaneously. Typically, there exist two kinds of approaches to cope with such constraints. One is to improve the capability of remote sensing instruments to make them competent for meeting the challenges of real-world applications; however, this is not only technically very difficult but also costly. The other is to blend or fuse distinct features retrieved from different data sources into one data set to enrich features of importance for complex real-world applications. Blending or fusing distinct features retrieved from different data sources turns out to be more practical for enhancing feature extraction.

Multiple observations with different characteristics from cross-mission sensors have provided unprecedented opportunities to collectively monitor the changing environment. Combining information from multiple data sources into a spatially and/or temporally regular dataset by using various modeling techniques to facilitate further data interpretation and analysis is considered a valuable tool and a common practice in remote sensing science. Regardless of whether these sensors were designed as a constellation in the beginning data fusion, or the process of blending information from multiple sources into a new data set, can be performed based on a variety of continuous data sources over the electromagnetic spectrum such as visible light, infrared light, and microwaves, as well as discrete data sources such as laser radar (e.g., LiDAR). If the input data sources for fusion are provided in terms of remotely sensed imagery, it can also be referred to as image fusion.

The basic concepts and relevant technology hubs of data fusion have been well elaborated in Chapter 8 through a thorough literature review.

Since the incipient stage of data fusion, a myriad of techniques and algorithms have been developed through various applications to advance the knowledge of data fusion. Due to different spatial, temporal, and spectral characteristics, simply using one algorithm to fuse all data sets is by no means a feasible solution. In principle, data fusion methods can be grouped into two classical categories. One is the spatiotemporal domain fusion, and the other is the transform domain fusion, each of which involves complex mathematical constructs to be used as instruments to model real-world events and to serve as an object of reasoning. The former may include Intensity-Hue-Saturation (IHS) transformation, Principal Component Analysis (PCA), and wavelet transforms; the latter can be further classified into different categories, including artificial intelligence-based fusion algorithms, probabilistic fusion algorithms, statistical fusion algorithms, and unmixing-based fusion. Some hybrid approaches exist to integrate the advantages from different data fusion methods. To improve the application potential, these recent data fusion algorithms may be grouped by function such as: (1) the enhancement of spatial resolution of multispectral and/or hyperspectral satellite imagery for terrestrial feature extraction (Chang et al., 2015), (2) the improvement of temporal resolution of satellite imagery for public health assessment (e.g., Liu and Weng, 2012), (3) the generation of advanced daily land surface temperature using extreme learning machine at Landsat resolution for climate change studies (Huang et al., 2013; Weng et al., 2014; Bai et al., 2015), (4) the optical fusion of RapidEye and MODIS imageries to explore vegetation dynamics of semi-arid rangelands (Tewes et al., 2015), (5) the fusion of multiple temporal, angular, and spectral features to improve land cover classification accuracy (e.g., Chen et al., 2017a), and (6) the exploration of the spatiotemporal distributions of environmental quality (Puttaswamy et al., 2013; Chang et al., 2014a,b; Chang et al., 2015; Xu et al., 2015a; Chang et al., 2016). Thus, the fused data or image should have enhanced spatial, temporal and/or spectral characteristics compared to each individual data set prior to the fusion process.

9.2 DATA FUSION TECHNIQUES AND ALGORITHMS

Conceptually, data fusion can be treated as a procedure that combines information from multiple data sources through a mathematical modeling framework (Figure 9.1), and the whole process can be further modeled as

$$O = F(X_1', X_2', \ldots, X_n') \tag{9.1}$$

where $(X_1', X_2', \ldots, X_n')$ denotes the transformed information from the original data files $(X_1, X_2, \ldots, X_n)$, and F represents a modeling framework to fuse the transformed information toward a final output O (Dai and Khorram, 1999). Specifically, F refers to various algorithms and techniques aiming to

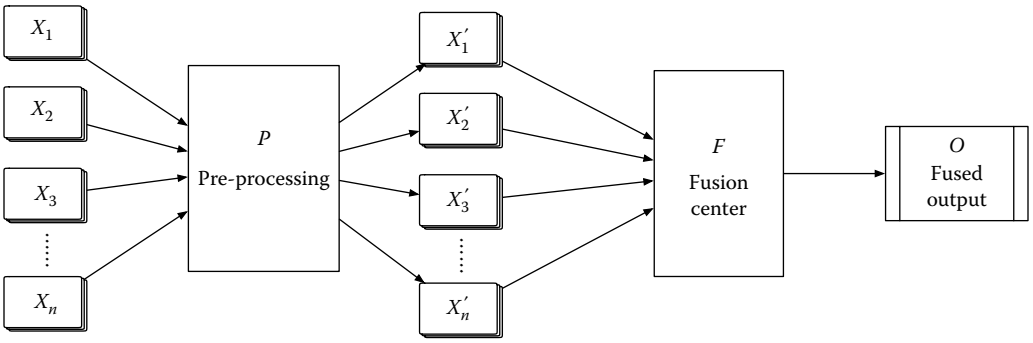

FIGURE 9.1 A general framework of data fusion.

blend information. Hence, it covers a large variety of methods. As previously discussed, image and data fusion methods can be divided into various categories, despite their fundamental principles and theories. In this section, these fundamental theories and principles lying behind each typical data fusion algorithm will be briefly introduced in addition to descriptions of their practical implementations in real-world environmental monitoring applications.

9.2.1 Pan-Sharpening

In the early stage of satellite remote sensing, in order to balance constraints between spatial, temporal, and spectral resolutions of remote sensing sensors as well as to keep satellites' weight, cost, bandwidth, and complexity down, a few sensors with limited capabilities were deployed onboard satellites. To enhance the Earth observing capabilities, satellite sensors providing high resolution panchromatic (PAN) images are commonly deployed in addition to sensors offering common medium/low resolution multispectral (MS) images. Configuring satellite sensors in such a way is a relatively common practice. For example, the well-known Landsat 7 incorporates sensors providing six 30-m resolution multispectral bands and a 60-m thermal infrared band together with a 15-m resolution panchromatic band. Such band combinations are commonly bundled in satellite remote sensing, and there exist a variety of similar satellite sensors onboard Earth observation platforms such as SPOT, QuickBird, OrbView, IKONOS, WorldView, and GeoEye, which commonly include both coarse spatial resolution MS bands and a single fine spatial resolution PAN band.

Pan-sharpening, shorthand for panchromatic sharpening, is a pixel-level fusion technique aimed at fusing a coarser spatial resolution MS image with a finer spatial resolution PAN image to create a new dataset that has the spectral properties of MS with the spatial resolution of PAN images. It aims to increase the spatial resolution of the MS to match that of the PAN images, while preserving the initial spectral information simultaneously (Amro et al., 2011; Vivone et al., 2015). In other words, the PAN image is used to sharpen the MS image; "to sharpen" means "to improve" the spatial resolution of the MS image. A generic pan-sharpening process typically encompasses the following procedures, including (1) up-sampling and interpolation of MS images, (2) alignment and co-registration between MS and PAN images, and (3) the core fusion steps (Figure 9.2).

With the increasing availability of high-resolution remotely sensed images, the demand for pan-sharpened data is continuously growing. As an important method to enhance images by improving the spatial resolution of the MS images, pan-sharpening enables us to perform a large

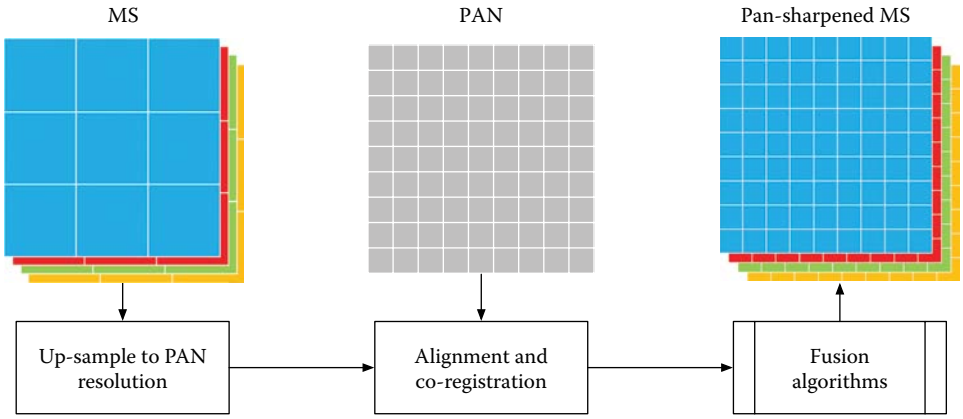

FIGURE 9.2 A schematic illustration of a generic pan-sharpening procedure. MS and PAN denote multispectral and panchromatic images, respectively.

variety of remote sensing tasks that require images with both high spatial and spectral resolutions, such as simple visual image interpretation (Laporterie-Déjean et al., 2005), change detection (Souza et al., 2003), object recognition and classification (Mohammadzadeh and Zoej, 2006; Aguilar et al., 2013), and higher-level product generation and advanced manipulation (Günlü et al., 2014).

A myriad of image fusion algorithms has been developed for meeting different pan-sharpening objectives over the past decades. In the public domain, there exists an enormous amount of pan-sharpening techniques and methods, which even date back to the 1980s. For example, the Intensity Modulation (IM) integration method was proposed for the integration of Landsat Multispectral Scanner (MSS) and Seasat Synthetic Aperture Radar (SAR) images by modulating the intensity of each pixel of the MSS channels with the value of the corresponding pixel of the SAR image (Wong and Orth, 1980). The actual pan-sharpening emerged after the launch of the French SPOT satellite system in February of 1986 when PAN images became available. Meanwhile, the paradigm term for image fusion shifted from the original "integration" to "merge" by taking advantage of IHS transformation (Welch and Ehlers, 1987; Tu et al., 2004). Since then, a variety of methods have been proposed for sharpening MS images with PAN data, such as High Pass Filtering (HPF) (Chavez and Bowell, 1988), Brovey Transform (BT) (Chavez Jr. et al., 1991; Bovolo et al., 2010), PCA (González-Audícana et al., 2004; Yang et al., 2007, 2012b) and Gram-Schmidt (GS) method (Laben and Brower, 2000; Aiazzi et al., 2006; Dalla Mura et al., 2015). In addition to these classical methods, pan-sharpening can also be performed by making use of methods like wavelet transform (Zhou et al., 1998), contourlet transform (Amro and Mateos, 2010), and generalized Laplacian pyramid (Aiazzi et al., 2002), as well as statistical methods such as spatially adaptive fusion methods (Aiazzi et al., 2002), Bayesian-based approaches (Fasbender et al., 2008), and many others (Wang et al., 2014).

Previous studies regarding various pan-sharpening methods attempted to classify them into distinct categories based on principles used in the fusion process. For example, Ranchin and Wald (2000) classified them into three groups including projection method, substitution method, and relative spectral contribution method, and Schowengerdt (2006) further classified these methods briefly into spatial domain and spectral domain. In general, it is hard to find a universal classification of methods in the literature (Aiazzi et al., 2009; Amro et al., 2011). According to Amro et al. (2011), the pan-sharpening methods can be grouped into the following typical categories.

9.2.1.1 Component Substitution

The component substitution (CS) methods are classical approaches for pan-sharpening which work through projecting the MS image into another space, assuming that this CS output isolates the spatial structure from the spectral information. Subsequently, the transformed MS image can be enhanced by replacing the component that contains the spatial structure with the PAN image. Then the final pan-sharpened MS image can be generated by reversely transforming the component-substituted MS image back to the original color space (Shettigara, 1992; Thomas et al., 2008; Amro et al., 2011). Specifically, a CS method usually involves the following essential steps (Figure 9.3): (1) up-sample the MS image to the same spatial resolution as the PAN, (2) transform the up-sampled MS image into a set of components, (3) match the histogram of the PAN with one of the transformed components of MS (i.e., aiming to reduce possible distortion as the greater the correlation between the PAN image and the replaced component, the lower the distortion introduced by the data fusion process), (4) replace that component with the histogram matched PAN image, and (5) reversely transform all components (with the substituted) back to the original color space to obtain a pan-sharpened MS image. Methods that work in such a way to sharpen MS images (Figure 9.4a) include the commonly used IHS (Carper et al., 1990), PCA (Chavez and Kwarteng, 1989) (Figure 9.4b), GS (Farebrother, 1974; Laben and Brower, 2000) (Figure 9.4c), and BT (Chavez Jr. et al., 1991; Bovolo et al., 2010) (Figure 9.4d), because all these methods involve performing a different transformation

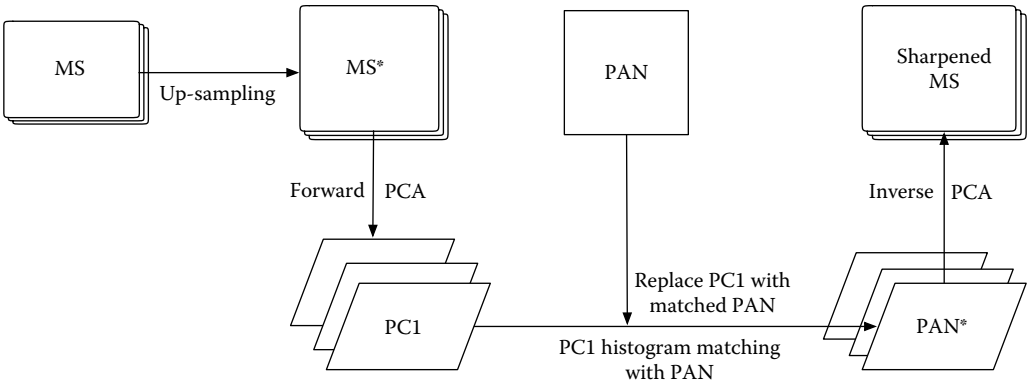

FIGURE 9.3 Pan-sharpening with PCA transform.

on the MS image. Among these methods, the GS algorithm was found to excel over other techniques due to its better performance in spectral preservation and spatial injection (Aiazzi et al., 2009).

Compared to other methods, CS techniques are global, meaning they operate in the same way on the whole image. Hence, techniques belonging to this category enable us to render the spatial details into the final product (Aiazzi et al., 2007). In general, CS techniques are fast and easy

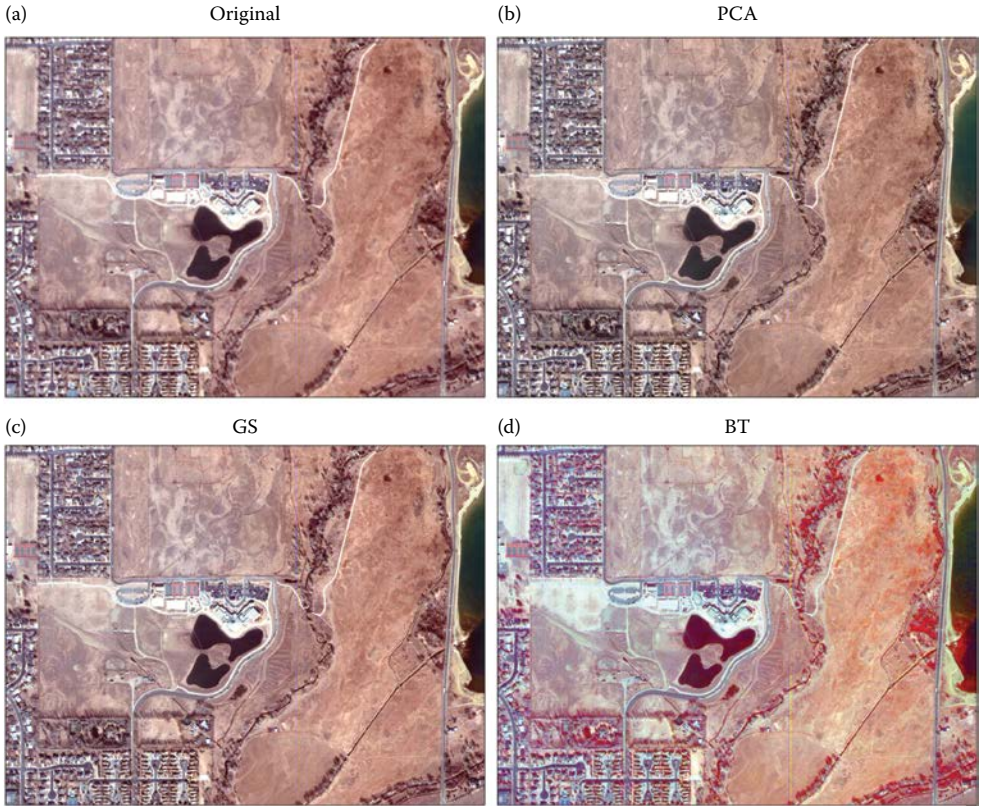

FIGURE 9.4 Pan-sharpened images using different fusion techniques. (a) One scene of Multispectral Imaging (MSI) image over Boulder, Colorado, USA, (b) pan-sharpened with PCA method, (c) pan-sharpened with GS method, and (d) pan-sharpened with BT method. Both the MSI (with a spatial resolution of 2.8 m) and PAN (with a spatial resolution of 0.7 m) images were acquired by instruments onboard the QuickBird satellite.

to implement and provide sharpened images with good visual/geometrical quality in most cases (Aiazzi et al., 2007). On the other hand, the global operating capability makes CS techniques incapable of accounting for local dissimilarities between the PAN and MS images originating from the spectral mismatch between the PAN and MS channels of the instruments, which may produce significant spectral distortions (Thomas et al., 2008). In order to find one component that is most suitable for substitution, statistical tests or weighted measures are incorporated to adaptively select an optimal component for substitution and transformation, and such methods are known as adaptive component substitution (Shah et al., 2008; Rahmani et al., 2010). A more detailed description of the main CS methods can be found in Vivone et al. (2015).

9.2.1.2 Relative Spectral Contribution

The relative spectral contribution (RSC) family can be considered a variant of the CS family, as a linear combination of the spectral bands is used instead of substitution. The process involves (1) up-sampling the MS image to the same spatial resolution as the PAN image, (2) matching the histogram of the PAN with each up-sampled MS band, and (3) pan-sharpening each up-sampled MS band by a linear combination of the spectral information at each pixel (*i, j*) through the following mathematical computations:

$$HRMS_b(i,j) = \frac{MS_b^h(i,j) \times PAN(i,j)}{\sum_b MS_b^h(i,j)} \tag{9.2}$$

where $HRMS_b$ denotes the pan-sharpened MS image and MS_b^h represents the high resolution (up-sampled) MS image at band b ($b = 1,2,\ldots,B$).

Methods falling in the RSC family include BT and IM. BT is a simple method of merging data from different sensors based on the chromaticity transform, aiming at increasing the visual contrast in the high and low ends of the data's histogram, with an assumption that the spectral range spanned by the PAN image is the same as that being covered by the multispectral channels (Gillespie et al., 1987). Because only three bands are involved, a pan-sharpened image with BT has great spectral distortions and hence is not suitable for pixel-based classification as the pixel values are changed drastically (Vijayaraj et al., 2004) (Figure 9.4d). The IM method can be performed by replacing the summation of all MS bands in Equation 9.2 with the intensity component of the IHS transformation (Vijayaraj et al., 2004). As reported in the literature, the IM method may result in color distortions if the spectral range of the intensity replacement (or modulation) image is different from the spectral range covered by the three bands used in the color composition (Alparone et al., 2004). Comparisons between CS and RSC methods have been discussed in detail in Tu et al. (2001) and Wang et al. (2005).

9.2.1.3 High Frequency Injection

The concept of high frequency injection (HFJ) is to transfer the high-frequency details extracted from the PAN image into the up-sampled MS images by applying spatial filtering techniques (Tsai, 2003). An early idea for such an approach was initialized by Chavez (1986), who aimed to extract the spectral information from the Landsat TM and combine it with the spatial information from a data set with a much higher spatial resolution. Relevant methods include HPF (Chavez and Bowell, 1988) and high pass modulation (HPM) (Schowengerdt, 2006), which work through the following steps: (1) un-sample the MS image according to the resolution of the PAN image, (2) perform low pass filtering on the PAN image, (3) derive a high frequency image by subtracting the filtered image from the original PAN, and (4) obtain a pan-sharpened image by adding the high frequency image to each band of the MS. Further investigations indicate that the HFJ methods have superior performance compared with many other pan-sharpening methods such as those in the CS family due to the HFJ methods' lower spectral distortion and edge noises (Wald et al., 1997).

9.2.1.4 Multi-Resolution Transformation

Spatial transformation methods like convolution and Fourier transform have long been considered a valuable tool for extracting the spatial information of remote sensing images. In order to access spatial information over a wide range of scales or directions from local to global in addition to time and frequency, methods such as Laplacian pyramids (Burt and Adelson, 1983; Aiazzi et al., 2002), wavelet transform (Zhou et al., 1998), contourlet transform (Amro and Mateos, 2010), and curvelets transform (Dong et al., 2015a) can also be applied as a data representation beside these two extreme transformations. The basic assumption of such methods is that the activity level of source images can be measured by the decomposed coefficients in a selected transform domain (Liu et al., 2017a). Among these methods, wavelet transform has been widely used in pan-sharpening fusion schemes, which provides a framework to decompose images into a hierarchy with a decreasing degree of resolution, separating detailed spatial information between successive levels (Burt and Adelson, 1983). The essence of such a method is to decompose MS and PAN images into different components with variation scales so that spatial details can be enhanced in the sharpened MS images (Zhang, 2010).

Similar to CS and HFJ methods, multi-resolution transformation (MRT)-based data fusion can also be performed at the pixel level, but MRT methods enable the spectral adjustment of PAN images to be compatible with each MS band prior to the data fusion process while maintaining the original colors and statistical parameters (Zhou et al., 1998; Chang et al., 2016). Despite such advancements, MRT approaches may still exhibit certain spatial distortions (e.g., ringing artifacts) which decrease the visual quality of the fused data (Masi et al., 2016). The pan-sharpening process using wavelet and/or contourlet transform can be summarized as follows (Amro et al., 2011): (1) forward transform PAN and MS images using a sub-band and directional decomposition such as the subsampled or non-subsampled wavelet or contourlet transform, (2) apply a fusion rule onto the transform coefficients, and (3) obtain the pan-sharpened image by performing the inverse transform.

9.2.1.5 Statistical and Probabilistic Methods

In addition to the aforementioned methods, pan-sharpening can also be realized by making use of a set of methods that exploits the statistical characteristics of the MS and PAN images. By taking advantage of the substantial redundancy existing in the PAN data and the local correlation between PAN and MS images, Price (1987, 1999) successfully combined PAN and MS imageries from dual-resolution satellite instruments; the method was further improved by Park and Kang (2004) and named after the spatially adaptive algorithm. Compared to the original method, the spatially adaptive algorithm features adaptive insertion of information according to the local correlation between the two images, which sharpens the MS images and prevents spectral distortion as much as possible. By using a nonlocal parameter optimization scheme, MS bands are also sharpened by evaluating band-dependent generalized intensities (Garzelli et al., 2008; Garzelli, 2015). Additionally, Bayesian framework has also been broadly used for pan-sharpening by considering the available prior knowledge about the expected characteristics of the pan-sharpened image, which renders the problem of pan-sharpening into a probabilistic framework (Fasbender et al., 2008). These statistical methods have also been extensively used in fusing other types of remotely sensed imagery in addition to pan-sharpening; more details related to such data fusion methods will be introduced in the following subsections.

In recent years, several new kinds of pan-sharpening methods have been developed, emerging as a new branch in this field (e.g., Toet, 1989; Zhang and Guo, 2009; Yang and Li, 2010; Liang et al., 2012; Liu et al., 2017b). To date, there exists a large variety of methods and algorithms available for pan-sharpening, and comparisons of these methods have been well discussed in the literature (Chavez et al., 1991; Wang et al., 2005; Ehlers et al., 2010; Amro et al., 2011; Vivone et al., 2015). In most cases, pan-sharpening methods were applied to fuse MS and PAN images for the enhancement of the spatial details of MS while preserving the spectral characteristics. Nevertheless, with the

availability of hyperspectral images, pan-sharpening has been also exploited for data fusion of PAN and hyperspectral data (Licciardi et al., 2012; Loncan et al., 2015). It is clear that conventional methods are not suitable for this task due to possible issues such as nonsimultaneous acquisition, co-registration of the data, and different spatial coverages and resolutions (Vivone et al., 2015).

9.2.2 Statistical Fusion Methods

As previously discussed, the core process of data fusion can be treated as one statistical modeling procedure, aiming to compute a new data value for each pixel by performing certain mathematical operations on the given input. In such a context, the data fusion based on CS methods can be considered one special case, as substitutions are performed rather than complex mathematical computations. On the other hand, although the missing spatial details of the low-resolution MS image can be enhanced with the aid of high-resolution PAN images simply by making use of methods such as CS and MRT, possible spectral distortions are introduced into the fused image if the PAN image is not exactly equivalent to the structural component of the low-resolution MS image (Zhong et al., 2016). In many cases, the data fusion can be also performed based on two MS images and/or hyperspectral images rather than simply based on MS and PAN images. More advanced methods should be used to model the relevant relations between information from each single sensor to account for data heterogeneities and spatial dependencies, in particular the feature extraction for multitemporal change detection over different terrestrial and aquatic environments (Chang et al., 2014b; Doña et al., 2015; Imen et al., 2015).

Generally, data fusion methods taking advantage of any statistical framework to characterize the inherent relationships between different features of the input images can all be referred to as statistical data fusion. Methods falling in this category always involve a statistical modeling process that plays an essential and critical role in addressing spatial, temporal, and spectral heterogeneities arising from instrumental, observational, and algorithmic differences between distinct inputs. To date, a large number of statistical modeling algorithms have been developed and used to fuse various remote sensing data/images in real-world applications. In general, these approaches can be broadly grouped into the following categories.

9.2.2.1 Regression-Based Techniques

Regression approaches have been used to fuse multiple datasets synergistically by constructing a regressed model between the target output and various data inputs, especially for cases with diversified data sources. For instance, by constructing a set of multiple linear regression models, the spatiotemporal distribution of mud content was clearly mapped to monitor the sediment grain-size of intertidal flats in the Westerschelde of the Netherlands through the data fusion of information from both space-borne microwave (SAR) and optical/shortwave infrared remote sensing data (van der Wal and Herman, 2007). Similarly, Srivastava et al. (2013) applied a multiple linear regression approach to fuse soil moisture data from the satellite soil moisture dedicated mission (SMOS) and WRF-NOAH Land Surface Model (WRF-NOAH LSM) to improve the accuracy of soil moisture deficit estimation. Based on a regression tree analysis, Kellndorfer et al. (2010) succeeded in fusing various remote sensing data (from LiDAR, InSAR, and Landsat) for forest stand height characterization by modeling the relationship between LiDAR measured canopy height and a suite of potential predictor variables derived from InSAR and optical datasets. Data fusion based on such methods, to some extent, can be considered a convolution of multiple data sources toward a synergistic output with improved accuracy or enhanced features.

9.2.2.2 Geostatistical Approaches

As the remotely sensed data are often provided in terms of 2-dimensional imagery with detailed spatial information associated, spatial relationships between various images can be modeled and incorporated to aid in data fusion. Differing from other types of data fusion methods, geostatistical

solutions provide another family of statistical fusion approaches by explicitly incorporating the concept of spatial analysis (e.g., spatial correlation). Kriging, one classic geostatistical approach that is commonly used in spatial interpolation applications, has been also explored as a potential tool for image fusion objectives (Pardo-Igúzquiza et al., 2006; Atkinson et al., 2008; Meng et al., 2010; Sales et al., 2013). The theoretical basis of Kriging was proposed by Georges Matheron in 1960, and the basic idea is to predict the value of a function at a given location by computing a weighted average of the known values of the function near that point. To a larger extent, Kriging can be also considered a kind of regression analysis based on a single random field rather than multiple observations. To date, a variety of Kriging estimators has been developed in addition to the simple one, including Ordinary Kriging, Universal Kriging, Indicator Kriging, Disjunctive Kriging, Co-Kriging, and Kriging with External Drift. Descriptions of such estimators have been detailed in Mitchell (2007).

Due to the significant advantage of preserving spectral properties of the observed coarse resolution images, Kriging-based methods have been widely used in image fusion applications for downscaling purposes (Nishii et al., 1996; Pardo-Igúzquiza et al., 2006; Chatterjee et al., 2010; Meng et al., 2010; Nguyen et al., 2012; Puttaswamy et al., 2013; Sales et al., 2013). For instance, with the aid of a traditional universal Kriging approach and a spatial statistical data fusion scheme, Puttaswamy et al. (2013) successfully merged two satellite-based daily Aerosol Optical Depth (AOD) measurements with ground-based observations over the continental United States to improve spatial coverage and overall accuracy. As summarized in Sales et al. (2013), the primary advantages of Kriging over other techniques are mainly due to its capacity in accounting for: (1) pixel size differences between two inputs with different spatial resolutions by considering the sensor's point spread function, (2) the spatial correlation structure of the attribute values at image pixels, (3) the preservation of the spectral characteristics of the imagery, and (4) the utilization of all spectral information in the data fusion procedure.

In recent years, a set of more advanced Kriging methods was successfully developed and applied for various image downscaling and/or pan-sharpening purposes. For instance, Meng et al. (2010) developed an image fusion scheme based on Regression Kriging (RK) to fuse multitemporal high-resolution satellite images by taking consideration of the correlation between response variables (i.e., the image to be fused) and predictor variables (i.e., the image with finer spatial resolution), spatial autocorrelation among pixels in the predictor images, and the unbiased estimation with minimized variance. In 2013, Sales et al. (2013) proposed a method of Kriging with External Drift (KED) to downscale five 500-m MODIS pixel bands to match two 250-m pixel bands. Later in 2015, a new method termed Area-To-Point Regression Kriging (ATPRK) was proposed, also aiming to downscale 500-m MODIS bands 3–7 to a resolution of 250-m (Wang et al., 2015a). The ATPRK method involves two essential steps: (1) regression modeling aiming to incorporate fine spatial resolution ancillary data, and (2) Area-To-Point Kriging (ATPK)-based residual downscaling (i.e., downscaling coarse residuals to the desired fine spatial resolution) (Figure 9.5). ATPRK enables the preservation of the spectral properties of the original coarse data and the extension of the spectral properties to other supplementary data (Wang et al., 2015a, 2017a). Based on ATPRK, two more extended methods, termed adaptive ATPRK (Wang et al., 2016) and spectral–spatial adaptive ATPRK (Zhang et al., 2017), were proposed for pan-sharpening and MODIS image downscaling purposes, respectively, with improvements focusing mainly on enhancing spectral fidelity and spatial details in the data fusion process.

9.2.2.3 Spatiotemporal Modeling Algorithms

In most cases, the input data for data fusion are always provided in terms of remotely sensed images at different spatial and temporal scales, for instance, fine temporal resolution but coarse spatial resolution MODIS images and coarse temporal resolution but fine spatial resolution Landsat data. In order to generate data with both fine spatial and temporal resolutions, the relevant spatial and temporal differences must be well modeled and accounted for in the data fusion process. Hence,

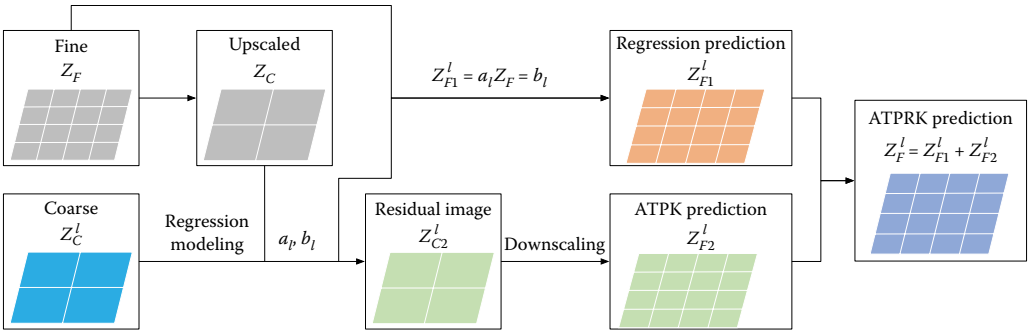

FIGURE 9.5 A schematic illustration of the use of the ATPRK algorithm to fuse coarse and fine images. (Adapted from Wang, Q., Shi, W., and Atkinson, P. M., 2016. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 151–165.)

techniques and algorithms aimed at modeling spatial and temporal relationships between the input data sources for meeting fusion objectives are often categorized as Spatio-Temporal Data Fusion (STDF) methods. In recent years, a large number of STDF methods have been developed under different assumptions and application contexts. Among them, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) is the most well-known (Gao et al., 2006). STARFM is a typical image-pair based approach that was originally developed with the intention of fusing daily MODIS land surface reflectance (1,000-m) with 16-day Landsat TM reflectance (30-m) to generate synthetic Landsat-like imagery with a spatial resolution of 30-m on a daily basis (Figure 9.6).

Typically, STARFM involves the following four essential steps: (1) coarse MODIS images are first reprojected and resampled to the fine Landsat imagery in order to have the identical spatial resolution and georeference frame, (2) a moving window with fixed size is used to slide over the Landsat imagery to identify spectrally similar neighboring pixels, (3) an optimal weight is calculated for each neighbor pixel relying on spectral, temporal, and spatial differences, and (4) the surface reflectance of the central pixel of the moving window is computed in a modeling framework. The key is to model spatiotemporal variations between the up-sampled coarse image and the observed fine image to account for the relevant spectral and temporal differences. More specifically, STARFM works by relying on at least one coarse-fine image pair on temporally close days as well as one

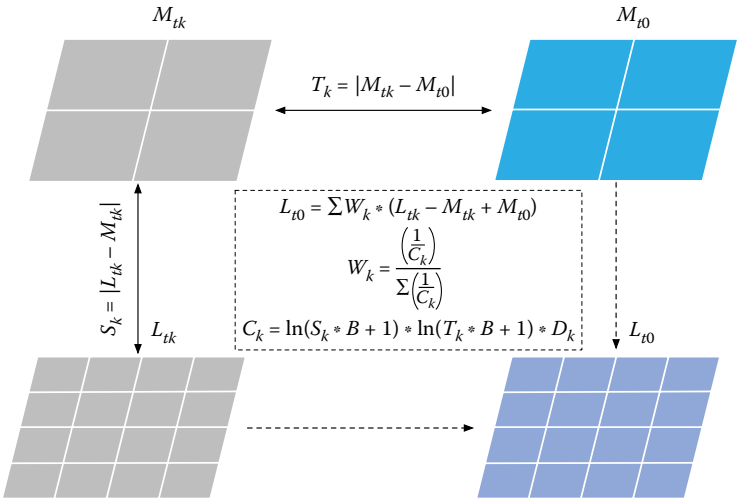

FIGURE 9.6 A schematic illustration of the STARFM algorithm based on one coarse-coarse and one coarse-fine image pair for data fusion.

coarse image to obtain a fused image with fine spatial resolution on the prediction day (e.g., Figure 9.6). Such a method assumes that pixels of the same class or adjacent pixels should have similar pixel values, and hence the temporal variations of the high spatial resolution image in one moment can be predicted by matching the neighboring similar pixels at different temporal phases.

Mathematically, the modeling framework can be considered a linearly weighted combination of the coarse temporal variations added to the available fine spatial resolution image, as the weights are determined by spectral difference, temporal difference, and location distance (Wang et al., 2017b). Regarding Figure 9.6, the prediction process can be modeled as (Gao et al., 2006):

$$L(x_{w/2}, y_{w/2}, t0) = \sum_{i=1}^{w}\sum_{j=1}^{w} W_{ij}(L(x_i, y_j, tk) - M(x_i, y_j, tk) + M(x_i, y_j, t0)) \tag{9.3}$$

where $L(x_{w/2}, y_{w/2}, t0)$ is a predicted pixel value in Landsat-like imagery for the time $t0$. w denotes the size of the moving window (e.g., a default of 1,500 m in STARFM) and hence $(x_{w/2}, y_{w/2})$ denotes the central pixel in this moving window. $L(x_i, y_j, tk)$ and $M(x_i, y_j, tk)$ are the observed Landsat and up-sampled MODIS reflectance values at time tk, which are used as base images to characterize spectral differences between two distinct sensors. $M(x_i, y_j, t0)$ is the up-sampled MODIS reflectance value at the window location (x_i, y_j) observed at $t0$. As shown in Equation 9.3, the core of the modeling is to determine weight W_{ij} for each neighboring pixel (x_i, y_j), which is calculated as a normalized reverse distance of combined weighting function C_{ij}.

$$W_{ij} = (1/C_{ij}) / \sum_{i=1}^{w}\sum_{j=1}^{w} (1/C_{ij}) \tag{9.4}$$

where C_{ij} denotes the combined weighting calculated based on spectral (S_{ij}) and temporal (T_{ij}) differences as well as spatial distance (D_{ij}),

$$C_{ij} = S_{ij} * T_{ij} * D_{ij} \tag{9.5}$$

$$S_{ij} = \left| L(x_i, y_j, tk) - M(x_i, y_j, tk) \right| \tag{9.6}$$

$$T_{ij} = \left| M(x_i, y_j, tk) - M(x_i, y_j, t0) \right| \tag{9.7}$$

$$D_{ij} = 1.0 + \frac{\sqrt{(x_{w/2} - x_i)^2 + (y_{w/2} - y_j)^2}}{A} \tag{9.8}$$

where A is a constant that defines the relative importance of spatial distance to the spectral and temporal distance; hence, a smaller value of A yields a larger dynamic range of D_{ij} (Gao et al., 2006). It is clear that STARFM works by fully relying on modeling statistical relations between image pairs at the pixel level, without performing complex transformation or substitution.

Due to its adaptive nature, STARFM has been extensively used in practice to aid in environmental surveillance by providing fused images with enhanced spatial, temporal, and spectral characteristics, such as land cover change detection (Hilker et al., 2009b; Walker et al., 2012), evapotranspiration mapping (Cammalleri et al., 2014), flood mapping (Zhang et al., 2014), water quality mapping (Chang et al., 2014b; Doña et al., 2015; Imen et al., 2015), and public health studies (Liu and Weng, 2012). Despite the salient capacity in generating fused data with both high spatial resolution and frequency coverage, it should be noted that there exist several limitations associated with STARFM,

as reported in the literature (Hilker et al., 2009a; Zhu et al., 2010). First, STARFM is inadequate for predicting abrupt land cover changes that are not recorded in any baseline fine resolution image. This is mainly due to the fact that the predicted Landsat-like image is a linear weighted combination of the baseline fine resolution image and the modeled temporal changes. Second, STARFM fails to handle directional dependence of reflectance as the method simply treats the input data as a purely statistical number. Last, STARFM does not perform well on heterogeneous landscapes because temporal changes are derived from pure, homogeneous patches of land cover at the MODIS pixel scale.

In order to solve these limitations, a set of STARFM-like or STARFM-based methods was proposed in recent years. In 2009, Hilker et al. (2009a) developed a Spatial and Temporal Adaptive Algorithm for Mapping Reflectance CHange (STAARCH) to fuse Landsat and MODIS reflectance data for forest disturbance mapping purposes. By taking advantage of the tasseled cap transform of two or more image pairs, STAARCH enables the detection of spatial and temporal changes with a high level of detail and thus improves the final accuracy of the synthetic images. In 2010, an Enhanced STARFM (ESTARFM) was developed by making use of the observed reflectance trend between two points in time and spectral unmixing theory, aiming to improve the final prediction accuracy in heterogeneous landscapes (Zhu et al., 2010). A simple performance comparison of these three methods can be found in Gao et al. (2015). To address the spatial autocorrelation issue embedded in ESTARFM, Fu et al. (2013) modified a similar pixel selection scheme and thus derived a new method termed ESTARFM.

Aside from the methods discussed above, many other available techniques may also fall into this category due to their statistical nature. Random weighting, for instance, was used to fuse multisensor observations by adopting the weights of each individual sensor to obtain a final optimal weight distribution (Gao et al., 2011). Similarly, Bisquert et al. (2015) proposed a simple and fast data fusion scheme based on a weighted average of two input images (i.e., one coarse-fine image pair), considering their temporal validity to the image to be fused. Recently, dictionary-pair learning was also applied to STDF (e.g., Huang and Song, 2012; Song and Huang, 2013; Wei et al., 2016). Huang and Song (2012) developed a Sparse-Representation-based Spatiotemporal Reflectance Fusion Model (SPSTFM) to model temporal changes between coarse-fine image pairs, with the trained dictionary used to predict the unknown fine spatial resolution reflectance based on the modeled temporal changes among coarse images. Such a method was further extended to perform image fusion simply based on one image pair (Song and Huang, 2013; Chen et al., 2017b). In addition to remote sensing applications, statistical fusion methods are also a valuable tool in the biomedical domain. For example, by using a convex optimization method known as semidefinite programming (Vandenberghe and Boyd, 1996), a set of kernel-based heterogeneous descriptions of the same set of genes was fused together to provide better visualization and understanding of the gene structure (Lanckriet et al., 2004).

9.2.3 Unmixing-Based Fusion Methods

The unmixing approach commonly uses a linear mixing model to analyze mixed pixels, which assumes that the spectrum of a mixed pixel is a linear combination of the pure spectra of the components present in that pixel weighted by their fractional coverage (Settle and Drake, 1993; Zurita-Milla et al., 2009). Mathematically, the spectral property $C(t,b)$ of a pixel with coarse resolution (known as a mixed pixel) at spectral band b at date t can be modeled as a weighted summation of the mean spectral properties of different endmembers (with a total number of n) derived from a fine resolution image $\overline{F(c,t,b)}$ and their abundances $A(c,t)$, regardless of errors arising from other external impacts such as atmospheric effects.

$$C(t,b) = \sum_{c=1}^{n} \overline{F(c,t,b)}\, A(c,t) + \varepsilon \tag{9.9}$$

If there exists *a priori* knowledge about both the components that might be present in a given scene and their pure spectra, their sub-pixel proportions can then be retrieved through a linear mixing model in a process known as spectral unmixing (Adams et al., 1995). Due to its sound physical basis and effectiveness in analyzing mixed pixels, spectral unmixing has been widely used in the remote sensing community for various applications, such as the quantification of land cover change at sub-pixel scales (Kresslerl and Steinnocher, 1999; Haertel et al., 2004). It is clear that the performance of spectral unmixing relies largely on the quality of the *a priori* knowledge with respect to the scene composition and its pure spectral quality. Hence, spectral unmixing is preferable for hyperspectral images, whereas it cannot be performed on panchromatic images (Zurita-Milla et al., 2009).

One of the most important tasks of image fusion is to improve the spatial resolution of coarse resolution images for spatial detail enhancements; this process is often referred to as downscaling. The linear mixing model can also be used to downscale the spectral information of the coarse image to the spatial resolution identical to the fine image, and this application is commonly known as spatial unmixing (Zhukov et al., 1999; Zurita-Milla et al., 2008, 2011). Spatial unmixing is different from spectral unmixing, however. The former aims to estimate the class endmembers within each coarse pixel based on the known class proportions that are calculated from the fine spatial resolution image, and the latter is designed to estimate the class proportions within the coarse pixel based on the class endmembers that are predetermined by either endmember extraction or reference to supervised information (Wang et al., 2017b). To be more specific, spatial unmixing does not require *a priori* knowledge of the main components present in the low spatial resolution image, as the input (pure signal) and spectral unmixing is in fact the output of spatial unmixing (Zurita-Milla et al., 2009). Therefore, spatial unmixing can even be applied with mixed pixels or a small number of spectral bands in coarse resolution images.

Figure 9.7 depicts a theoretical workflow of the unmixing-based data fusion algorithm, which primarily consists of the following four steps (Zhukov et al., 1999; Zurita-Milla et al., 2008, 2011; Gevaert and García-Haro, 2015; Zhu et al., 2016): (1) define endmembers at coarse resolution by clustering the input fine resolution data, (2) calculate endmember proportions of each coarse pixel, (3) unmix the coarse pixels within a given moving window based on the computed endmember fractions in the previous step, and (4) assign the unmixed spectral value to fine pixels. There exists

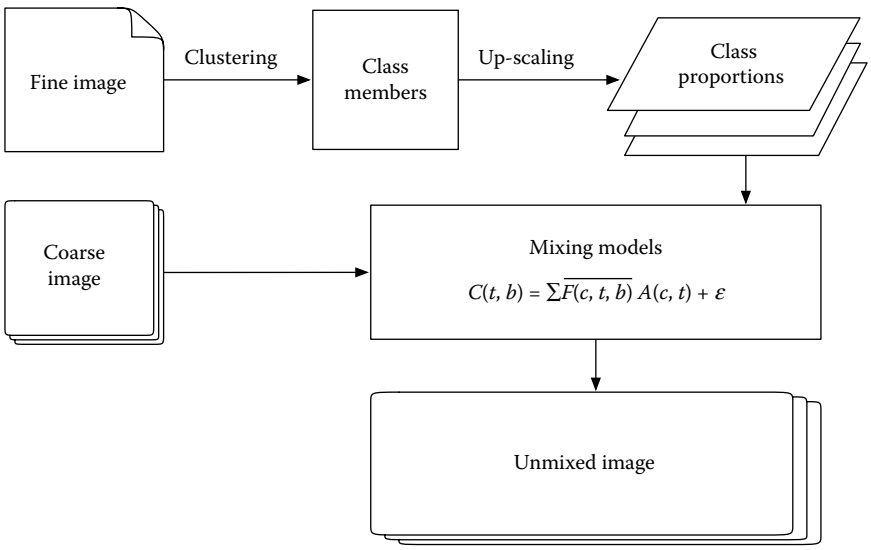

FIGURE 9.7 A schematic illustration of spatial unmixing-based fusion algorithm.

a big assumption associated with spatial unmixing methods, that is, no land-cover/land-use changes occur during the period as the endmember fractions remain constant for each coarse image (Wang et al., 2017b).

An early framework of spatial unmixing was initialized by Zhukov et al. (1999), by developing a multisensor, multi-resolution technique to fuse low- and high-resolution images acquired at different times for a synergetic interpretation. This method was further enhanced by performing various optimizations to the processes prior to unmixing in order to gain possible accuracy improvements via: (1) accounting for both spatial distance and spectral dissimilarity between pixels to tune weights assigned to the linear mixing model (Busetto et al., 2008), (2) optimizing the number of endmembers and the size of the moving window used to solve the unmixing equations (Zurita-Milla et al., 2008), and (3) making use of a high spatial resolution land use database (Zurita-Milla et al., 2009).

Recently, Amorós-López et al. (2011) developed a regularized spatial unmixing method to downscale MEdium Resolution Imaging Spectrometer (MERIS) full resolution images by making use of Landsat/TM images; they added a regularization term to the cost function of the spatial unmixing to limit large deviations of the unmixed pixels for each class. This method was later used for crop monitoring by generating a series of Landsat-like MERIS images (Amorós-López et al., 2013), and was further improved to downscale MODIS data by incorporating prior class spectra (Xu et al., 2015b). In 2012, Wu et al. (2012) proposed a Spatial and Temporal Data Fusion Model (STDFM) to generate Landsat-like surface reflectance from MODIS by considering both the spatial variation and the nonlinear temporal change information. In STDFM, the prediction was estimated through unmixing endmember reflectance at both input and prediction date and then adding the estimated change back to the base fine-resolution image under the assumption that the temporal variation properties of each land-cover class are constant (Wu et al., 2012). That is,

$$F(t0,b) = F(tk,b) + \Delta U(\Delta tk,b) \tag{9.10}$$

where

$$\Delta U(\Delta tk,b) = U(t0,b) - U(tk,b) \tag{9.11}$$

Here, $F(t0, b)$ denotes the predicted fine spatial resolution image for band b at time $t0$ based on one base image at time tk. U is the unmixed image and ΔU represents the variation between two unmixed images.

To fuse coarse resolution hyperspectral and fine resolution multispectral data, Yokoya et al. (2012) developed a method termed coupled nonnegative matrix factorization unmixing, with the main principle of combining abundance matrices from both hyperspectral and multispectral data after alternately applying unmixing to them. Later in 2013, an enhanced version of STDFM (ESTDFM) was proposed that incorporated a patch-based ISODATA classification method as well as making use of a sliding window to better determine the endmember in order to better fuse Landsat and MODIS surface reflectance (Zhang et al., 2013). In 2015, Gevaert and García-Haro (2015) proposed a method called the Spatial and Temporal Reflectance Unmixing Model (STRUM) which incorporated the main features of STARFM in order to obtain temporally stable synthetic imagery at Landsat spatial resolution. Similarly, by integrating spectral unmixing analysis with a thin plate spline interpolator, Zhu et al. (2016) developed a method termed Flexible Spatiotemporal DAta Fusion (FSDAF) for fusing satellite images with different resolutions. The FSDAF method features minimum input data requirements, heterogeneous landscapes accommodations, and the capability to predict both gradual change and land cover type change (Zhu et al., 2016).

9.2.4 Probabilistic Fusion Methods

As claimed by Zhang and Huang (2015), all the CS and MRT methods can be derived from a Bayesian fusion framework by adjusting a weight parameter to balance contributions from the spatial injection and spectral preservation models. This claim strongly indicates the strength and superiority of the Bayesian approach in advancing the data fusion modeling process. Here, data fusion methods taking advantage of probabilistic theories like Bayesian are all referred to as probabilistic fusion methods. In addition to the traditional Bayesian method (Solberg et al., 1994; Mascarenhas et al., 1996; Fasbender et al., 2008, 2009), some other well-known methods such as maximum likelihood (Xu et al., 2015a), Markova random field (Nishii, 2003), random forest (Liu et al., 2014), and Dempster-Shafer evidence theory (Bendjebbour et al., 2001; Lu et al., 2015) that work in a probabilistic framework have been widely used in data fusion processes. A great advantage of performing data fusion using probability theory is its ability to model the problem in a probabilistic framework without restricting modeling hypotheses (Fasbender et al., 2008).

In data fusion frameworks, the Bayesian theory enables us to take into account the uncertainties in the input multi-scale data. The straightforward formulation of such a data fusion method is often fast and easy to implement. In a Bayesian fusion paradigm, Bayesian Maximum Entropy (BME), one of the nonlinear spatiotemporal geostatistical methods that can theoretically blend multi-scale data from different sources with different accuracies (Li et al., 2013b), has been widely used to fuse related data from multiple sources in a probabilistic manner with a non-Gaussian distribution hypothesis (Christakos and Li, 1998). A basic workflow to fuse multiple data making use of BME can be found in Li et al. (2013a,b) and Tang et al. (2016). Theoretically, in BME fusion framework, data with uncertainties (e.g., images at coarser resolution) are routinely considered soft (or fuzzy) data that can be expressed by probability distribution, whereas the accurate data (e.g., images at finer resolution) are treated as hard data with a probability of 1 (Li et al., 2013b). Hence, both types of data can be effectively integrated based on a Bayesian paradigm.

Mathematically, the BME process to blend multiple observations for a new estimation $\widehat{x_k}$ at the location (x, y) at time t can be expressed as

$$\widehat{x_k} = \int x_k f^*(x_k | x_{soft}, x_{hard}) d_{x_k} \tag{9.12}$$

where $f^*(x_k|x_{soft}, x_{hard})$ denotes a posterior Probability Density Function (PDF) over the spatiotemporal adjacent pixel observations (Li et al., 2013b). x_{soft} and x_{hard} are probabilistic soft data and hard data, respectively. Specifically, the hard data refer to those data obtained with high accuracy; for instance, ground-based instruments measured temperature at sparse monitoring stations. In contrast, the soft data refer to data sets with uncertainties, such as satellite retrievals, model simulated results, assimilated data sets, and so on. Because the posterior PDF at the estimation point is derived from the prior PDF in the Bayesian rule when soft data are involved, the posterior PDF can be written as

$$f^*(x_k | x_{soft}, x_{hard}) = \frac{f(x_{soft}, x_{hard}, x_k)}{f(x_{soft}, x_{hard})} = \frac{f_G(x_{map})}{f(x_{soft}, x_{hard})} \tag{9.13}$$

where $f(x_{soft},x_{hard})$ is *a priori* PDF at the spatiotemporal adjacent pixels and f denotes a detailed distribution function, for example, Gaussian distribution. $f_G(x_{map})$ is the joint PDF that can be attained by maximizing the entropy under the given constraint of the general knowledge G, which may consist of physical laws, scientific theories, logical principles, and summary statistics (Li et al., 2013b).

As elucidated in Li et al. (2013a) and Tang et al. (2016), the entropy H in BME framework can be defined as

$$H = -\int f_G(x_{map}) \log f_G(x_{map}) d_{x_{map}} \tag{9.14}$$

Introducing the Lagrange multipliers λ_α maximizes the entropy H equivalent to maximize the following relation:

$$\begin{aligned} L\left[f_G(x_{map})\right] = &-\int f_G(x_{map}) \log f_G(x_{map}) d_{x_{map}} \\ &-\sum_{\alpha}^{N} \lambda_\alpha \left[\varphi_\alpha(x_{map}) f_G(x_{map}) d_{x_{map}} - \overline{\varphi_\alpha(x_{map})}\right] \end{aligned} \tag{9.15}$$

where $L[f_G(x_{map})]$ is the objective function for maximizing entropy, $\varphi_\alpha(x_{map})$ is a set of functions x_{map} such as the mean and covariance moments, and $\overline{\varphi_\alpha(x_{map})}$ is the expected value of $\varphi_\alpha(x_{map})$. Based on Equation 9.15, the maximum entropy solution for $f_G(x_{map})$ can be attained by setting the partial derivatives to zero and solving the system of equations with respect to the λ_α

$$f_G(x_{map}) = \frac{\exp\left[\sum_{\alpha}^{N} \lambda_\alpha \varphi_\alpha(x_{map})\right]}{\int \exp\left[\sum_{\alpha}^{N} \lambda_\alpha \varphi_\alpha(x_{map})\right] d_{x_{map}}} \tag{9.16}$$

The final estimation of $\widehat{x_k}$ can be finally attained by replacing $f_G(x_{map})$ in Equations 9.12 and 9.13 (Tang et al., 2016).

Because of its effectiveness in accounting for differences between multiple data sources, BME has been widely used to integrate *a priori* knowledge and uncertain data to enrich the subjective information and then consider their uncertainties in deriving much more objective results. For example, Li et al. (2013b) succeeded in fusing thermal infrared and microwave sea surface temperature (SST) products to improve the accuracy, spatial resolution, and completeness of satellite SST products. A similar framework was also used to fuse leaf area index (Li et al., 2013a), chlorophyll-a concentration (Shi et al., 2015), and Aerosol Optical Depth (AOD) (Christakos and Li, 1998). Similarly, Fasbender et al. (2009) developed a space–time prediction scheme to account for secondary information sources contributing to the variation of nitrogen dioxide concentration based on a Bayesian data fusion framework. Such a fusion framework can also be applied for pan-sharpening (Fasbender et al., 2008). Likewise, by making use of maximum likelihood, Xu et al. (2015a) proposed a data fusion framework to fuse multiple AOD products aiming to create a consistent AOD dataset with improved spatial coverage.

Markov Random Field (MRF) is a set of random variables with a Markov property described by an undirected graph which has a representation of dependencies similar to Bayesian networks but with an undirected nature. Thus, to some extent, an MRF can represent certain dependencies that a Bayesian network cannot. Because of this capability, MRF has been widely used to model various low- to mid-level tasks in image processing and computer vision (Li, 2009), for example, data fusion. For instance, Xu et al. (2011) developed an MRF-based data fusion framework to fuse multi-spectral images by incorporating the contextual constraints via MRF models into the fusion model. Similarly, Sun et al. (2013) developed an MRF-based gradient domain image fusion framework, in which the salient structures of input images were fused in the gradient domain and then the

final fused image was reconstructed by solving a Poisson equation which forces the gradients of the fused image to be close to the fused gradients. In the framework, an MRF model was applied to accurately estimate region-based fusion weights for the salient objects or structures. Similarly, in order to fuse SAR images and optical imagery to estimate the nutrient fertility (total inorganic nitrogen) in coastal waters, Liu et al. (2014) developed a data fusion framework by making use of the random forest method. In the fusion framework, the random forest algorithm was used to merge four different input variables from the SAR and optical imagery to generate a new dataset (i.e., total inorganic nitrogen).

With the development of probabilistic theory, more advanced methods such as evidential belief reasoning can also be applied in probabilistic fusion framework. Dempster-Shafer (D-S) evidence theory, initiated in Dempster's work (Dempster, 1968) and mathematically formalized by Shafer (Shafer, 1976), has long been considered a generalization to the Bayesian theory and a popular tool to deal with uncertainty and imprecision (Khaleghi et al., 2013). Different from traditional Bayesian methods, D-S theory introduces the notion of assigning belief and plausibility to possible measurement hypotheses along with the required combination rule in order to fuse them (Khaleghi et al., 2013).

Mathematically, if we consider X to be all possible states of a system and 2^X to be all possible subsets of X, D-S theory will assign a belief mass m rather than a probability mass to each element E of 2^X to represent possible propositions with respect to the system state X. Here, the belief mass function m constrains the following two properties (Khaleghi et al., 2013):

$$m(\varnothing) = 0 \tag{9.17}$$

$$\sum_{E \in 2^X} m(E) = 1 \tag{9.18}$$

where $m(E)$ represents the level of evidence or confidence in E. If $m(E) > 0$, then the subset E of 2^X is referred to as a focal element (Zhou et al., 2013). Based on the mass function m, a probability range can be obtained for a given E

$$bel(E) \leq P(E) \leq pl(E) \tag{9.19}$$

Here, $bel(E)$ and $pl(E)$ refer to the belief of E and the plausibility of E, respectively, which can be calculated as

$$bel(E) = \sum_{B \in E} m(B) \tag{9.20}$$

$$pl(E) = \sum_{B \cap E \neq \varnothing} m(B) \tag{9.21}$$

If there exist two information sources with a belief mass function of m_1 and m_2, this information can then be fused based on Dempster's rule of combination (Khaleghi et al., 2013)

$$m_{1,2}(E) = (m_1 \oplus m_2)(E) = \frac{\sum_{B \cap C = E \neq \varnothing} m_1(B) m_2(C)}{1 - K} \tag{9.22}$$

where $m_{1,2}$ is the joint belief mass function, which should follow $m_{1,2}(\varnothing) = 0$. K represents the amount of conflict between the sources, which can be computed as

$$K = \sum_{B \cap C = \varnothing} m_1(B) m_2(C) \tag{9.23}$$

It is clear that D-S theory differs from traditional Bayesian inference by allowing each source to contribute information with different levels of detail. In other words, D-S theory is more flexible (Khaleghi et al., 2013). Due to this advantage, D-S theory has become a promising approach to fuse multisensor observations in recent years. For example, in order to better map and monitor woodland resources, Lu et al. (2015) proposed a D-S evidence theory-based fusion framework, aiming to fuse classification results from polarimetric synthetic aperture radar (PolSAR) (classified in a supervised manner) and interferometric SAR (InSAR) (classified in an unsupervised manner) data. The results indicated that the fused products had higher accuracy (95%) than that of each individual data set before fusion (88% and 90%). In general, probabilistic fusion methods could yield better performance compared to other simple statistical fusion approaches due to their abilities to reduce uncertainty in the fusion modeling process.

9.2.5 Neural Network-Based Fusion Methods

The critical step in data fusion process is modeling the relevant relations between various data inputs at different scales. Linear models have been widely used to establish relations between different data inputs because of their simplicities and effectiveness. Despite such advantages, it is noticeable there exist deficiencies related to the fact that many real-world systems are not consistent with such a linear assumption nonlinear. It is always difficult to obtain an accurate result by simply making use of linear models. For example, linear mixing models are often used to unmix coarse resolution images by referring to calculated endmember fractions from the fine resolution images, under an assumption that the spectral property of a coarse pixel is a linear weighted summation of various endmembers within this pixel. In other words, the relationship between them has already been prescribed to such a linear combination, which may in turn limit a full exploration toward higher prediction accuracy.

To cope with these deficiencies, nonlinear methods, such as artificial intelligence-based approaches, can be of great help. Among them, the most well-known approach is the Artificial Neural Network (ANN). ANNs represent a set of decision processing models inspired by the a biological neural network that is a series of interconnected neurons whose activation defines a recognizable linear pathway acting as decision aggregates to global decisions (Figure 9.8).

$$net_i = \sum_j \omega_{i,j} x_j + b_i \tag{9.24}$$

$$O_i = f(net_i) \tag{9.25}$$

where ω denotes the weights between two connected neurons that can be determined by various learning methods. x represents various inputs and b is the bias added to the weighted inputs. By applying an activation function f (e.g., typically a sigmoid function of $(1+\exp(-k))^{-1}$) to the net input, the output O can be obtained. Because of its nature, ANN enables the capture and representation of complex input/output relationships through a learning process without making assumptions regarding data distribution or the nature of the relation between inputs and outputs. Additionally, the learned empirical knowledge can be saved and further used for future data

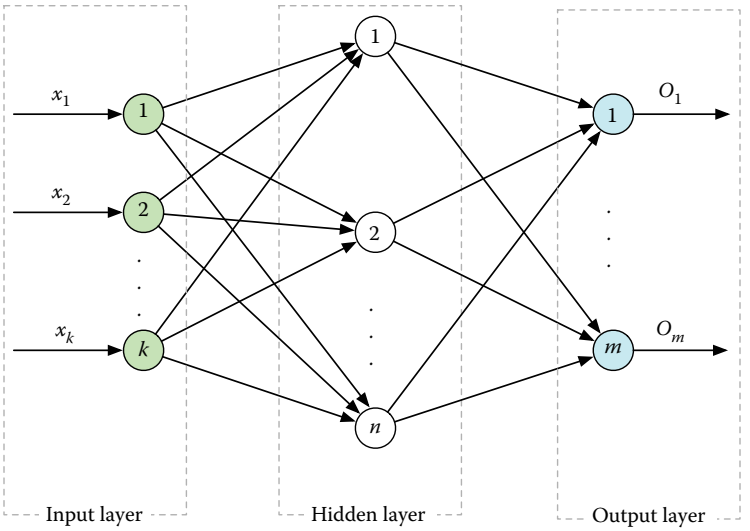

FIGURE 9.8 An example of a typical neural network architecture with one single hidden layer.

prediction purposes. As the learning process can be solved in a global optimization manner, the neural network approach is error tolerant and relatively insensitive to noise. All these features make the neural network approach a powerful tool with sound modeling capabilities, which is capable of generating solutions to complex problems beyond formal descriptions in spite of the drawback that neural networks are often considered black box models (Del Carmen Valdes and Inamura, 2001).

Because of the inherent distribution-free advantage, the neural network approach has long been considered a valuable tool for the advancement of data fusion techniques. Specifically, the neural network approach has the inherent ability to integrate information from multisource data and to identify complex relationships between mixed pixels (Dai and Khorram, 1999). This unique distribution-free feature makes the neural network approach excel among statistical approaches because it avoids exploring a predetermined relationship posed on the fusion of multisource data and thus is adaptable to numerous data inputs. At present, a large variety of neural-scale-based fusion approaches are available for data fusion.

By using a Back-Propagation Neural-network (BPN) and a Polynomial Neural Network (PNN), distinct features from infrared (IR) multispectral surveillance data were fused to provide excellent discrimination applicability between heavy and light space objects (Haberstroh and Kadar, 1993). The basic principle is to learn patterns embedded in each input, then data fusion is often performed on the feature- and/or decision-level. Similar data fusion frameworks taking advantage of BPN were also used to analyze multitemporal change of land cover characterization (Dai and Khorram, 1999), spatial resolution enhancement (Valdés Hernández and Inamura, 2000), and pan-sharpening (Del Carmen Valdes and Inamura, 2001). Apart from the conventional BPN, a set of more complex neural networks, such as Radial Basis Function (RBF)-based neural network (Chen and Li, 2006), pulse-coupled neural network (Li et al., 2006), Genetic Algorithm and Self-Organizing Feature Map (GA-SOFM) integrated artificial neural network (Yang et al., 2010), and MultiLayer Perceptron (MLP) neural network (Del Frate et al., 2014), has been applied to fuse multisensor, multifrequency (e.g., optical and microwave) remote sensing images in various practices.

With the recent development of theory in neural networks, several more advanced neural networks have been proposed and applied to advance data fusion practices. For instance, Huang et al. (2015) developed a Deep Neural Network (DNN)-based image fusion framework for pan-sharpening

purposes by fully taking advantage of the feature learning ability of DNN. The proposed DNN-based data fusion framework mainly consists of three procedures, including (1) patch extraction to generate the training set, (2) DNN training, and (3) final high-resolution image reconstruction. According to Huang et al. (2015), once a training set $\{x_p^i, y_p^i\}_{i=1}^N$ (N is the total number of training image patches) is extracted from the high-resolution PAN image ($\{x_p^i\}_{i=1}^N$) and the synthetic PAN image of the up-sampled low-resolution MS image ($\{y_p^i\}_{i=1}^N$), a DNN model can be trained by using the Modified Sparse Denoising Autoencoder (MSDA) algorithm to learn the relations between coarse and fine images. Mathematically, the feedforward functions of MSDA can be defined as:

$$h(y_p^i) = s(Wy_p^i + b) \tag{9.26}$$

$$\hat{x}(y_p^i) = s(W'h(y_p^i) + b') \tag{9.27}$$

where s is the sigmoid activation function, and W (or W') and b (or b') are the encoding (or decoding) weights and biases, respectively. $h(y_p^i)$ represents the hidden layer's activation, while $\hat{x}(y_p^i)$ denotes the reconstructed input that is an approximation of x_p^i. Parameters $\Theta = \{W, W', b, b'\}$ can then be trained by minimizing the following cost function with a Lagrangian approach:

$$L\left(\left\{x_p^i, y_p^i\right\}_{i=1}^N; \Theta\right) = \frac{1}{N}\sum_{i=1}^{N}\left\|x_p^i - \hat{x}\left(y_p^i\right)\right\|_2^2 + \frac{\lambda}{2}\left(\|W\|_F^2 + \|W'\|_F^2\right) + \beta KL(\hat{\rho}\,\|\,\rho) \tag{9.28}$$

where λ and β are two balancing parameters determined by cross validation. As indicated, the second term in the right is a weight decay term, and the last term $KL(\hat{\rho}\,\|\,\rho)$ is the Kullback–Leibler divergence, which is a sparsity term forcing $\hat{\rho}$ (the average activation function of the hidden layer) to approximate ρ:

$$KL(\hat{\rho}\,\|\,\rho) = \rho\log\frac{\rho}{\hat{\rho}} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}} \tag{9.29}$$

$$\hat{\rho} = \frac{1}{N}\sum_{i=1}^{N} h\left(y_p^i\right) \tag{9.30}$$

After training an MSDA, the successive layer is trained in turn by taking the hidden activation values $h(x_p^i)$ and $h(y_p^i)$ of the previous layer as the input in order to construct a deep learning network architecture; such a scheme is thus termed stacked MSDA (S-MSDA) (Huang et al., 2015). The experimental results show that such a data fusion scheme outperforms other pan-sharpening methods in terms of both visual perception and numerical measures. Similarly, Dong et al. (2015b) proposed a further deep learning method for single image super-resolution through an end-to-end mapping between the coarse- and fine-resolution images by making use of a deep Convolutional Neural Network (CNN). Compared to standard neural networks, CNN can yield better results with less training time because of its unique structures (Jing et al., 2017). Likewise, CNN can also be used for pan-sharpening (Masi et al., 2016; Zhong et al., 2016) as well as fusion of multispectral and hyperspectral images (Palsson et al., 2017).

Compared to other fusion methods, a neural-scale-based fusion scheme could yield higher accuracy for the fusion output, especially in dealing with complex data sources as no distribution priori are guaranteed. The reason can be largely attributed to the nonlinear, nonparametric, and distribution-free characteristics associated with neural scale methods. Despite the effectiveness of such methods, a possible disadvantage that should be noted is that the complex modeling process may require even more time than simple approaches such as linear regression or weighted summation.

9.2.6 Fuzzy Set Theory-Based Fusion Methods

Fuzzy set theory, initialized by Zadeh (1965) with the concept of partial membership to a set, has been found to provide a simple, yet powerful mathematic tool to model the vagueness and ambiguity in complex systems by imitating the human reasoning ability to reach conclusions with imprecise, incomplete, and not even totally reliable information (Solaiman et al., 1999; Melgani, 2004). Because of this distinct feature, fuzzy set theory has been used in a wide range of domains to process unprecise or uncertain data. In recent years, fuzzy-based data fusion methods were also developed and applied for various decision fusion schemes. For instance, Tupin et al. (1999) developed a fuzzy fusion approach to combine several structure detectors to facilitate SAR image classification. Likewise, a fuzzy-based multisensor data fusion classifier was developed for land cover classification by integrating multisensor and contextual information in a single and homogeneous framework (Solaiman et al., 1999). The integration was realized by iteratively updating the fuzzy membership maps corresponding to different thematic classes and by using spatial contextual membership information based on different fuzzy contextual classification rules. Similar works can also be found in the literature, for example, a fuzzy fusion of spectral, spatial, and temporal contextual information approaches for the classification of multisensor remote sensing images (Melgani, 2004), a fuzzy decision rule-based fusion for the classification of urban remote sensing images (Fauvel et al., 2006), a Choquet fuzzy integral-based fusion scheme to estimate the chlorophyll-a concentration in a water body (Wang et al., 2010), and a fuzzy K-nearest neighbor classification system to fuse hyperspectral and LiDAR data (Bigdeli et al., 2014a). Results from these studies indicate that the incorporation of fuzzy set theory in fusion framework could yield higher classification accuracy.

9.2.7 Support Vector Machine-Based Fusion Methods

In addition to the aforementioned approaches, accounting for uncertainties other high-level nonlinear, nonparametric modeling techniques can also be applied for data fusion purposes, such as SVM. SVM, well-known in the fields of machine learning and pattern recognition, works as a nonparametric classifier approach which aims to discriminate inputs into two separate classes by fitting an optimal linear separating hyperplane to the training samples in a multidimensional feature space (Vapnick, 1998). A detailed description of the general concept of SVM and a brief summary of SVM theory can be found in Burges (1998) and Waske and Benediktsson (2007), respectively. Waske and Benediktsson (2007) presented a decision fusion scheme to combine multisensor data with the aid of two SVMs, with one used for classification and the other for decision fusion. Different from previous decision fusion frameworks which simply fused the final classification outputs, the original outputs of each SVM discriminant function were fused instead. With a similar framework, hyperspectral and LiDAR data were successfully fused in the decision level by making use of a SVM-based classifier fusion system (Bigdeli et al., 2014b).

9.2.8 Evolutionary Algorithms

Essentially, the modeling framework of image/data fusion can be largely considered as an optimization problem, as the core is to find an optimal solution to the data fusion models so as to fuse multiple data sources seamlessly. Therefore, evolutionary algorithms can be used as promising approaches in helping solve complex data fusion problems. In the evolutionary computing domain, there exists a variety of evolutionary algorithms with different optimization strategies, such as Genetic Algorithm (GA) (Fraser, 1962), evolutionary programming (Fogel and Fogel, 1996), Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1995), and so forth. Among these algorithms, GA and PSO are two popular methods commonly used in the remote sensing domain to deal with various complicated optimization and search problems, thus both of them will be introduced in the following sections to delineate how they help in the image/data fusion process.

GA, one of the most well-known evolutionary algorithms, works based on the principle inspired by the mechanism of natural selection-more specifically, the biological genetics and evolution in Darwin's theory. GA has been demonstrated to possess a sound ability to solve many optimization and search problems due to its intelligent and nonlinear nature (Fraser, 1962). Generally, GA works in a manner of simulating "survival of the fittest" among individuals by evolving generation-by-generation to search for a better solution to an optimization problem (Melanie, 1998). In GA, the evolution starts from a population of randomly generated individuals through an iterative process, while each individual represents a possible solution to the problem that can be mutated and altered based on certain rules (e.g., selection, crossover, and mutation). In each iteration, the population of all individuals refers to a generation that is arranged in a structure analogous to chromosomes. To make the optimization process evolve at each step, GA will randomly select a set of individuals from the current population and use them as parents to produce children for the next generation. Although randomly selected, GA does not work in a purely random manner in practice. Instead, the randomness is constrained with the former generation information in order to direct the optimization into the search space with better performance. After evolving through a suite of generations, the population will finally evolve toward an optimal solution. More details with respect to the foundations of GA can be found in Rawlins (1991).

In general, GA works on a principle analogous to gene structure which behaves like chromosomes to evolve toward an optimal solution. Essentially, it is associated with the following foundations: (1) individuals within a population will compete with each other, (2) more offspring will be generated from those individuals winning the competition, (3) "good" genes (individuals) will propagate throughout the population to produce even better offspring (i.e., "survival of the fittest"), and (4) each successive generation will be more adaptable to the environment than the previous one. These unique properties fully resemble the process to search an optimal solution for a given optimization problem within a search space, as solutions with poor performance will be discarded in a stepwise manner. As opposed to the classical derivative-based optimization algorithm, GA has two distinct advantages: (1) GA generates a population of points rather than a single point at each iteration, and thus it is more efficient, and (2) GA uses a random generator to select the next population rather than a deterministic computation, which in turn renders robustness of the results.

Because of its capability and effectiveness in quickly and efficiently reaching a reasonable solution to a complex problem by searching through a large and complex search space, in image/data fusion domain, GA has been widely used in various fusion schemes to optimize either the fusion inputs or fused outputs toward the improvement of the final accuracy (Mumtaz et al., 2008; Lacewell et al., 2010). For instance, in order to maintain the detail and edge information from multisensor images, Li and Peng (2015) proposed a GA-aided cellular neural network based image fusion scheme, in which the GA algorithm was specifically applied for parameter optimization purposes. Analogously, a real-valued GA model was used to provide an optimum injection of PAN information into the MS images to improve the pan-sharpening performance (Garzelli and Nencini, 2006). Similar integrations can be also found in the literature, for example, GA and ANN (Peng and Dang, 2010), GA and morphological filters (Addesso et al., 2012), and so on.

The PSO algorithm was originally proposed in 1995 and was inspired by the flocking and schooling patterns of birds and fish (Kennedy and Eberhart, 1995). Analogous to GA, PSO is also a population-based optimization method that is commonly used to identify the optimal solution for evolutionary pathway problems. However, the PSO uses a different strategy as opposed to other algorithms to search for a similar optimal solution. In PSO, the optimization is performed based on a swarm of particles, and each particle represents a possible solution to the optimization problem in the search space. For each particle, there are two critical aspects associated with it, one is position and the other is velocity. Generally, the position of a particle is mainly influenced by the personal best position of this particle and the global best position (best position among the whole particles) in the swarm (Chen and Leou, 2012).

It is clear that each particle can evolve by learning from the previous experiences of all particles in the swarm, even the particle itself. During the evolving process, the position and velocity of each

particle are updated iteratively unless the optimum solution is attained while the potential solutions in each iteration will be updated as well until reaching a maximum number of iterations or a global optimum result (Kusetogullari et al., 2015). As claimed in Chen and Leou (2012), PSO is mainly featured by the following advantages as compared with other evolutionary algorithms: (1) Each particle in the swarm represents a potential solution that is "flying" within the hyperspace with a velocity. (2) Each particle can learn from the previous experiences of all particles as the swarm and particles have their own memories. This is the unique feature of PSO which does not exist in any other evolutionary algorithm. (3) Only primitive mathematical operations are required by PSO as the following states can be updated automatically.

As an effective off-line optimization method, PSO has been widely used in image/data fusion applications. For instance, Wang et al. (2015b) adopted a PSO algorithm to obtain the optimal weights that can be used for better pan-sharpening MS images by the PAN image, as the optimization was equivalent to the maximization of the radiometric similarities between the MS and the up-scaled PAN images. The results indicated that the PSO-aided fusion scheme showed superior performance to popular component substitution-based fusion methods (Wang et al., 2015b). Saeedi and Faez (2011) performed a similar work by integrating a multi-objective PSO algorithm and the shiftable contourlet transform for image pan-sharpening. Likewise, Raghavendra et al. (2011) successfully developed two distinct image fusion schemes by making use of PSO to find an optimal strategy to fuse sub-band coefficients of visible and near infrared images so as to improve the face verification performance. Many other applications can be also found in the literature, such as Siddiqui et al. (2011); Ahmed et al. (2013); Gharbia et al. (2016), and so on.

In addition to image/data fusion, PSO has been also used in remote sensing to deal with other optimization problems, including image classification (Ding and Chen, 2012; Soliman et al., 2012), feature extraction (dimensionality reduction) (Chang et al., 2014a; Yang et al., 2012a), image enhancement (super-resolution) (Erturk et al., 2014), change detection (Kusetogullari et al., 2015), and so on. In recent years, a set of extended PSO algorithms have been developed to advance the optimization process, such as adaptive PSO (Zhan et al., 2009), orthogonal PSO (Zhan et al., 2011), accelerated PSO (Yang et al., 2011), and so on.

9.2.9 Hybrid Methods

There is no doubt that each data fusion method has its own strengths and weaknesses. There is no single data fusion approach sufficient to fulfill all requirements, however; in other words, each method is more or less limited by some constraints (Chang et al., 2016). Hence, it is always advisable to integrate several data fusion techniques with different capabilities to cope with the relevant drawbacks from each individual method toward a more accurate or better result. Such an integrated data fusion framework is commonly referred to as hybrid data fusion. Because of the complementary strengths from various techniques, hybrid data fusion methods always yield better results than a single approach alone. Therefore, hybrid data fusion methods have been widely used in practice by remote sensing domains to solve various complex problems.

In order to enhance the classification accuracy of hyperspectral data, an integrated feature and decision level fusion framework was developed based on a set of algorithms, including projection pursuit, majority voting, max rule, min rule, and average rule (Jimenez et al., 1999). The experimental practices demonstrate that projection pursuit-based feature fusion and majority voting-based data fusion algorithms enable the extraction of more information from the training samples while avoiding the problem of overfitting and the Hughes phenomenon, and in turn the final classification accuracy is improved (Jimenez et al., 1999). Based on the concept of high pass filtering, González-Audícana et al. (2004) developed a new fusion scheme to fuse MS and PAN images based on the multi-resolution wavelet decomposition. In this fusion scheme, the improved IHS and PCA methods were both used to inject the spatial details of the PAN image into the MS one. Further comparisons indicate that the proposed hybrid fusion scheme yielded merged images with improved quality

with respect to those obtained by standard IHS, PCA, and standard wavelet-based fusion methods. Similarly, a hybrid wavelet-artificial intelligence fusion approach was developed to downscale MODIS land surface temporal data based on Landsat data in which the wavelets were used mainly to capture the properties of the signals in different directions (Moosavi et al., 2015).

Due to the powerful optimization capacity, GA has been widely used in data fusion schemes in conjunction with other fusion methods to improve the accuracy of fused outputs. In order to better retain the detail and edge information of the original images, a novel multisource image fusion scheme was developed based on cellular neural networks and GA, and the GA method was used to determine optimal parameters for the neural network fusion framework (Li and Peng, 2015). In addition to these complex methods, a hybrid fusion framework can also be attained by integrating several statistical approaches. For example, Wang et al. (2017b) developed an enhanced spatiotemporal data fusion framework aiming to increase prediction accuracy in fusing MODIS and Landsat images. To achieve this goal, two separate data fusion methods, ATPRK and STARFM, were used successively to fuse 250-m bands with 500-m bands of MODIS to produce the interim 250-m MODIS data while STARFM was used to blend the 250-m fused MODIS with 30-m Landsat images. Such a data fusion scheme enabled the capture of more information for abrupt changes and heterogeneous landscapes than simply using the original 500-m MODIS data, thus in turn increasing the accuracy of spatiotemporal fusion predictions (Wang et al., 2017b).

Although hybrid approaches could yield better performance in image fusion practices compared to using a single approach alone, comprehensive knowledge and understanding of each method are always required to guarantee accurate modeling and integration of these separate methods in the data fusion framework; otherwise, an improper integration could result in even more erroneous results. Generally, the degree of success for data fusion is often case-dependent. Nevertheless, there exist several niches in performing data fusion in order to better serve the users' needs (Pohl and Van Genderen, 1998). In an attempt to confirm such niches, there are a few issues that deserve our attention.

First, users should clearly know the primary objective of data fusion. More specifically, they should know the features that are to be enhanced. For instance, in meteorological applications, users may require data having frequent and repetitive coverage with relatively low spatial resolution; in contrast, users aiming to perform land cover change detection or urban mapping may demand an image with a higher spatial resolution, and military surveillance users may need to have images with both high spatial and temporal resolutions.

The second issue with which users should be familiar is related to the selection of the data for data fusion. At present, an enormous amount of satellite-based data is available, but with different characteristics which depend largely on the sensor capability, orbit of the platform, imaging geometry, ultimate resolution, and even bias levels and atmospheric constraints such as cloud cover and sun angle (Ardeshir Goshtasby and Nikolov, 2007; Nguyen et al., 2009; Castanedo, 2013). Thus, these factors need to be carefully considered in selecting proper data selection for the data fusion process, because a good data set can make the data fusion more successful. Third, the users also have to take into account the need for data pre-processing like up-sampling and registration, depending on which processing strategy one needs to follow.

Furthermore, the users should determine which technique to use for the problems at hand to explore the full advantage of the data available as different methods may yield distinct outputs. Last but not least, proper quality evaluation measures or criteria are necessary to assess the accuracy of the data fusion process. After data fusion, the accuracy of the fused outputs should be well assessed to guarantee a quality assured result. A set of statistical measures that can be used to assess the performance of data fusion results will be introduced in the next chapter.

9.3 SUMMARY

In this chapter, a set of common image and data fusion algorithms was introduced by providing details related to the underlying principles and theories of each method. STARFM, an advanced

multisensor data fusion algorithm originally proposed to create Landsat-like MODIS imagery by enhancing the spatial resolution of MODIS data, was elaborated here. In addition to STARFM, data fusion techniques making use of pan-sharpening, neural-network-based theory, probability theory, evolutionary algorithms, statistical approach, image unmixing, and many others were described as well. In general, data fusion approaches can be applied not only to enhance spatial, temporal, and/or spectral resolution of the input remotely sensed imagery, but also to improve prediction accuracy and to optimize data quality and spatial coverage simultaneously.

REFERENCES

Adams, J. B., Sabol, D. E., Kapos, V., Filho, R. A., Roberts, D., Smith, M. O., and Gillespie, A. R., 1995. Classification of multispectral images based on fractions and endmembers: Applications to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, 52, 137–152.

Addesso, P., Conte, R., Longo, M., Restaino, R., and Vivone, G., 2012. A pansharpening algorithm based on genetic optimization of Morphological Filters. In: *2012 IEEE International Geoscience and Remote Sensing Symposium*, 5438–5441.

Aguilar, M. A., Saldaña, M. M., and Aguilar, F. J., 2013. GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments. *International Journal of Remote Sensing*, 34, 2583–2606.

Ahmed, T., Singh, D., Gupta, S., and Raman, B., 2013. Partice swarm optimization based fusion of MODIS and PALSAR images for hotspot detection. In: *2013 IEEE International Conference on Microwave and Photonics (ICMAP)*, 1–6.

Aiazzi, B., Baronti, S., Alparone, L., and Garzelli, A., 2002. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40, 2300–2312.

Aiazzi, B., Baronti, S., Lotti, F., and Selva, M., 2009. A comparison between global and context-adaptive pansharpening of multispectral images. *IEEE Geoscience and Remote Sensing Letters*, 6, 302–306.

Aiazzi, B., Baronti, S., and Selva, M., 2007. Improving component substitution pansharpening through multivariate regression of MS+Pan data. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 3230–3239.

Aiazzi, B., Baronti, S., Selva, M., and Alparone, L., 2006. Enhanced Gram-Schmidt spectral sharpening based on multivariate regression of MS and pan data. In: *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 3806–3809.

Alparone, L., Facheris, L., Marta, S., Baronti, S., Cnr, I., Panciatichi, V., Garzelli, A., and Nencini, F., 2004. Fusion of multispectral and SAR images by intensity modulation. In: *Procceedings 7th International Conference on Information Fusion*, 637–643.

Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Moreno, J., and Camps-Valls, G., 2011. Regularized multiresolution spatial unmixing for ENVISAT/MERIS and landsat/TM image fusion. *IEEE Geoscience and Remote Sensing Letters*, 8, 844–848.

Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., and Camps-Valls, G., 2013. Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 23, 132–141.

Amro, I. and Mateos, J., 2010. Multispectral image pansharpening based on the contourlet transform. In: *Information Optics and Photonics: Algorithms, Systems, and Applications*, 247–261.

Amro, I., Mateos, J., Vega, M., Molina, R., and Katsaggelos, A. K., 2011. A survey of classical methods and new trends in pansharpening of multispectral images. *EURASIP Journal on Advances in Signal Processing*, 2011, 79.

Ardeshir Goshtasby, A., and Nikolov, S., 2007. Image fusion: Advances in the state of the art. *Information Fusion*, 8, 114–118.

Atkinson, P. M., Pardo-Igúzquiza, E., and Chica-Olmo, M., 2008. Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 573–580.

Bai, Y., Wong, M., Shi, W.-Z., Wu, L.-X., and Qin, K., 2015. Advancing of land surface temperature retrieval using extreme learning machine and spatio-temporal adaptive data fusion algorithm. *Remote Sensing*, 7, 4424–4441.

Bendjebbour, A., Delignon, Y., Fouque, L., Samson, V., and Pieczynski, W., 2001. Multisensor image segmentation using Dempster-Shafer fusion in Markov fields context. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1789–1798.

Bigdeli, B., Samadzadegan, F., and Reinartz, P., 2014a. Feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral and LIDAR data. *Journal of Applied Remote Sensing*, 8, 83509.

Bigdeli, B., Samadzadegan, F., and Reinartz, P., 2014b. A decision fusion method based on multiple support vector machine system for fusion of hyperspectral and LIDAR data. *International Journal of Image and Data Fusion*, 5, 196–209.

Bisquert, M., Bordogna, G., Bégué, A., Candiani, G., Teisseire, M., and Poncelet, P., 2015. A Simple fusion method for image time series based on the estimation of image temporal validity. *Remote Sensing*, 7, 704–724.

Bovolo, F., Bruzzone, L., Capobianco, L., Garzelli, A., Marchesi, S., and Nencini, F., 2010. Analysis of the effects of pansharpening in change detection on VHR images. *IEEE Geoscience and Remote Sensing Letters*, 7, 53–57.

Burges, J. C., 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121–167.

Burt, P. J. and Adelson, E. H., 1983. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31, 532–540.

Busetto, L., Meroni, M., and Colombo, R., 2008. Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series. *Remote Sensing of Environment*, 112, 118–131.

Cammalleri, C., Anderson, M. C., Gao, F., Hain, C. R., and Kustas, W. P., 2014. Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agricultural & Forest Meteorology*, 186, 1–11.

Carper, W. J., Lillesand, T. M., and Kiefer, R. W., 1990. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, 56, 459–467.

Castanedo, F., 2013. A review of data fusion techniques. *Science World Journal*, 2013, 1–19.

Chang, N.-B., Bai, K., Imen, S., Chen, C.-F., and Gao, W., 2016. Multisensor satellite image fusion and networking for all-weather environmental monitoring. *IEEE Systems Journal*, 1–17.

Chang, Y.-L., Liu, J.-N., Chen, Y.-L., Chang, W.-Y., Hsieh, T.-J., and Huang, B., 2014a. Hyperspectral band selection based on parallel particle swarm optimization and impurity function band prioritization schemes. *Journal of Applied Remote Sensing*, 8, 84798.

Chang, N.-B., Vannah, B. W., Yang, Y. J., and Elovitz, M., 2014b. Integrated data fusion and mining techniques for monitoring total organic carbon concentrations in a lake. *International Journal of Remote Sensing*, 35, 1064–1093.

Chang, Y.-L., Wang, Y. C., Fu, Y.-S., Han, C.-C., Chanussot, J., and Huang, B., 2015. Multisource Data Fusion and Fisher Criterion-Based Nearest Feature Space Approach to Landslide Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 576–588.

Chatterjee, A., Michalak, A. M., Kahn, R. a., Paradise, S. R., Braverman, A. J., and Miller, C. E., 2010. A geostatistical data fusion technique for merging remote sensing and ground-based observations of aerosol optical thickness. *Journal of Geophysical Research*, 115, D20207.

Chavez, P., 1986. Digital merging of Landsat TM and digitized NHAP data for 1: 24,000- scale image mapping. *Photogrammetric Engineering and Remote Sensing*, 52–10, 1637–1646.

Chavez, P. S. and Bowell, J. A., 1988. Comparison of the spectral information content of Landsat Thematic Mapper and SPOT for three different sites in the Phoenix, Arizona region. *Photogrammetric Engineering and Remote Sensing*, 54, 1699–1708.

Chavez, P. S. and Kwarteng, A. Y., 1989. Extracting spectral contrast in Landsat thematic mapper image data using selective principal component analysis. *Photogrammetric Engineering and Remote Sensing*, 55, 339–348.

Chavez Jr., P. S., Sides, S. C., and Anderson, J. A., 1991. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Photogrammetric Engineering and Remote Sensing*, 57, 295–303.

Chen, B., Huang, B., and Xu, B., 2017a. Multi-source remotely sensed data fusion for improving land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 124, 27–39.

Chen, B., Huang, B., and Xu, B., 2017b. A hierarchical spatiotemporal adaptive fusion model using one image pair. *International Journal of Digital Earth*, 10, 639–655.

Chen, H.-Y. and Leou, J.-J., 2012. Multispectral and multiresolution image fusion using particle swarm optimization. *Multimedia Tools and Applications*, 60, 495–518.

Chen, Y. W. and Li, B. Y., 2006. Remote sensing image fusion based on adaptive RBF neural network. In: *International Conference on Neural Information Processing*, 314–323.

Chen, S., Su, H., Zhang, R., Tian, J., and Yang, L., 2008. The tradeoff analysis for remote sensing image fusion using expanded spectral angle mapper. *Sensors*, 8, 520–528.

Christakos, G. and Li, X., 1998. Bayesian maximum entropy analysis and mapping: A farewell to kriging estimators? *Mathematical Geology*, 30, 435–462.

Dai, X. and Khorram, S., 1999. Data fusion using artificial neural networks: A case study on multitemporal change analysis. *Computers, Environment and Urban Systems*, 23, 19–31.

Dalla Mura, M., Vivone, G., Restaino, R., Addesso, P., and Chanussot, J., 2015. Global and local gram-schmidt methods for hyperspectral pansharpening. In: *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 37–40.

Del Carmen Valdes, M., and Inamura, M., 2001. Improvement of remotely sensed low spatial resolution images by back-propagated neural networks using data fusion techniques. *International Journal of Remote Sensing*, 22, 629–642.

Del Frate, F., Latini, D., Picchiani, M., Schiavon, G., and Vittucci, C., 2014. A neural network architecture combining VHR SAR and multispectral data for precision farming in viticulture. In: *2014 IEEE Geoscience and Remote Sensing Symposium*, 1508–1511.

Dempster, A. P., 1968. A generalization of bayesian inference. *Journal of the Royal Statistical Society Series B*, 30, 205–247.

Ding, S. and Chen, L., 2012. Spectral and spatial feature classification of hyperspectral images based on particle swarm optimisation. *International Journal of Innovative Computing and Applications*, 4, 233–242.

Doña, C., Chang, N.-B., Caselles, V., Sánchez, J. M., Camacho, A., Delegido, J., and Vannah, B. W., 2015. Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain. *The Journal of Environmental Management*, 151, 416–426.

Dong, C., Loy, C. C., He, K., and Tang, X., 2015b. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 295–307.

Dong, L., Yang, Q., Wu, H., Xiao, H., and Xu, M., 2015a. High quality multi-spectral and panchromatic image fusion technologies based on curvelet transform. *Neurocomputing* 159, 268–274.

Ehlers, M., Klonus, S., Johan Åstrand, P., and Rosso, P., 2010. Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion*, 1, 25–45.

Erturk, A., Gullu, M. K., Cesmeci, D., Gercek, D., and Erturk, S., 2014. Spatial resolution enhancement of hyperspectral images using unmixing and binary particle swarm optimization. *IEEE Geoscience and Remote Sensing Letters*, 11, 2100–2104.

Farebrother, R. W., 1974. Algorithm AS 79: Gram-schmidt regression. *Applied Statistics*, 23, 470.

Fasbender, D., Brasseur, O., and Bogaert, P., 2009. Bayesian data fusion for space-time prediction of air pollutants: The case of NO_2 in Belgium. *Atmospheric Environment*, 43, 4632–4645.

Fasbender, D., Radoux, J., and Bogaert, P., 2008. Bayesian data fusion for adaptable image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1847–1857.

Fauvel, M., Chanussot, J., and Benediktsson, J. A., 2006. Decision fusion for the classification of urban remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 2828–2838.

Fogel, D. B. and Fogel, L. J., 1996. An Introduction to Evolutionary Programming, In: *Artificial Evolution. AE 1995. Lecture Notes in Computer Science*, 21–33.

Fraser, A. S., 1962. Simulation of genetic systems. *Journal of Theoretical Biology*, 2, 329–346.

Fu, D., Chen, B., Wang, J., Zhu, X., and Hilker, T., 2013. An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model. *Remote Sensing*, 5, 6346–6360.

Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., and Yang, Y., 2015. Fusing landsat and MODIS data for vegetation monitoring. *IEEE Geoscience and Remote Sensing Magazine*, 3, 47–60.

Gao, F., Masek, J., Schwaller, M., and Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 2207–2218.

Gao, S., Zhong, Y., and Li, W., 2011. Random weighting method for multisensor data fusion. *IEEE Sensors Journal*, 11, 1955–1961.

Garzelli, A., 2015. Pansharpening of multispectral images based on nonlocal parameter optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 53, 2096–2107.

Garzelli, A. and Nencini, F., 2006. PAN-sharpening of very high resolution multispectral images using genetic algorithms. *International Journal of Remote Sensing*, 27, 3273–3292.

Garzelli, A., Nencini, F., and Capobianco, L., 2008. Optimal MMSE pan sharpening of very high resolution multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 228–236.

Gevaert, C. M. and García-Haro, F. J., 2015. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment*, 156, 34–44.

Gharbia, R., Baz, A. H. El, and Hassanien, A. E., 2016. An adaptive image fusion rule for remote sensing images based on the particle swarm optimization. In: *2016 IEEE International Conference on Computing, Communication and Automation (ICCCA)*, 1080–1085.

Gillespie, A. R., Kahle, A. B., and Walker, R. E., 1987. Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques. *Remote Sensing of Environment*, 22, 343–365.

González-Audícana, M., Saleta, J. L. J. L., Catalán, R. G., García, R., Gonzalez-Audicana, M., Saleta, J. L. J. L., Catalan, R. G., and Garcia, R., 2004. Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1291–1299.

Günlü, A., Ercanlı, İ., Sönmez, T., and Başkent, E. Z., 2014. Prediction of some stand parameters using pan-sharpened IKONOS satellite image. *European Journal of Remote Sensing*, 47, 329–342.

Haberstroh, R. and Kadar, I., 1993. Multispectral data fusion using neural networks. In: *1955, Signal Processing, Sensor Fusion, and Target Recognition II*, 65–75.

Haertel, V., Shimabukuro, Y. E., and Almeida, R., 2004. Fraction images in multitemporal change detection. *International Journal of Remote Sensing*, 25, 5473–5489.

Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., Gao, F., and White, J. C., 2009a. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*, 113, 1613–1627.

Hilker, T., Wulder, M. A., Coops, N. C., Seitz, N., White, J. C., Gao, F., Masek, J. G., and Stenhouse, G., 2009b. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing of Environment*, 113, 1988–1999.

Huang, B. and Song, H., 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 3707–3716.

Huang, B., Wang, J., Song, H., Fu, D., and Wong, K., 2013. Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring. *IEEE Geoscience and Remote Sensing Letters*, 10, 1011–1015.

Huang, W., Xiao, L., Wei, Z., Liu, H., and Tang, S., 2015. A new pan-sharpening method with deep neural networks. *IEEE Geoscience and Remote Sensing Letters*, 12, 1037–1041.

Imen, S., Chang, N.-B., and Yang, Y. J., 2015. Developing the remote sensing-based early warning system for monitoring TSS concentrations in Lake Mead. *Journal of Environmental Management*, 160, 73–89.

Jimenez, L. O., Morales-Morell, A., and Creus, A., 1999. Classification of hyperdimensional data based on feature and decision fusion approaches using projection pursuit, majority voting, and neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1360–1366.

Jing, L., Wang, T., Zhao, M., and Wang, P., 2017. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors*, 17, 414.

Kellndorfer, J. M., Walker, W. S., Lapoint, E., Kirsch, K., Bishop, J., and Fiske, G., 2010. Statistical fusion of lidar, InSAR, and optical remote sensing data for forest stand height characterization: A regional-scale method based on LVIS, SRTM, Landsat ETM+, and ancillary data sets. *Journal of Geophysical Research: Biogeosciences* 115, 1–10.

Kennedy, J. and Eberhart, R., 1995. Particle swarm optimization. In: *Proceedings of ICNN'95-International Conference on Neural Networks*. 1942–1948.

Khaleghi, B., Khamis, A., Karray, F. O., and Razavi, S. N., 2013. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14, 28–44.

Kresslerl, F. P. and Steinnocher, K. T., 1999. Detecting land cover changes from NOAA-AVHRR data by using spectral mixture analysis. *International Journal of Applied Earth Observation and Geoinformation*, 1, 21–26.

Kusetogullari, H., Yavariabdi, A., and Celik, T., 2015. Unsupervised change detection in multitemporal multispectral satellite images using parallel particle swarm optimization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 2151–2164.

Laben, C. and Brower, B., 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. United States Patent, 6.

Lacewell, C. W., Gebril, M., Buaba, R., and Homaifar, A., 2010. Optimization of image fusion using genetic algorithms and discrete wavelet transform. In: *Proceedings of the IEEE 2010 National Aerospace & Electronics Conference*, 116–121.

Lanckriet, G. R. G., De Bie, T., Cristianini, N., Jordan, M. I., and Noble, W. S., 2004. A statistical framework for genomic data fusion. *Bioinformatics*, 20, 2626–2635.

Laporterie-Déjean, F., de Boissezon, H., Flouzat, G., and Lefèvre-Fonollosa, M.-J., 2005. Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images. *Information Fusion*, 6, 193–212.

Li, S. Z., 2009. *Markov Random Field Modeling in Image Analysis*. Computer Science Workbench. Springer-Verlag, London, UK.

Li, A., Bo, Y., and Chen, L., 2013a. Bayesian maximum entropy data fusion of field-observed leaf area index (LAI) and Landsat Enhanced Thematic Mapper Plus-derived LAI. *International Journal of Remote Sensing*, 34, 227–246.

Li, A., Bo, Y., Zhu, Y., Guo, P., Bi, J., and He, Y., 2013b. Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method. *Remote Sensing of Environment*, 135, 52–63.

Li, M., Cai, W., and Tan, Z., 2006. A region-based multi-sensor image fusion scheme using pulse-coupled neural network. *Pattern Recognition Letters*, 27, 1948–1956.

Li, J. and Peng, Z., 2015. Multi-source image fusion algorithm based on cellular neural networks with genetic algorithm. *Optik-International Journal for Light and Electron Optics*, 126, 5230–5236.

Liang, J., He, Y., Liu, D., and Zeng, X., 2012. Image fusion using higher order singular value decomposition. *IEEE Trans. Image processing*, 21, 2898–2909.

Licciardi, G. A., Khan, M. M., Chanussot, J., Montanvert, A., Condat, L., Jutten, C., 2012. Fusion of hyperspectral and panchromatic images using multiresolution analysis and nonlinear PCA band reduction. *EURASIP Journal on Advances in Signal Processing*, 2012, 207.

Liu, Z., Chai, Y., Yin, H., Zhou, J., and Zhu, Z., 2017a. A novel multi-focus image fusion approach based on image decomposition. *Information Fusion*, 35, 102–116.

Liu, Y., Chen, X., Peng, H., and Wang, Z., 2017b. Multi-focus image fusion with a deep convolutional neural network. *Information Fusion*, 36, 191–207.

Liu, M., Liu, X., Li, J., Ding, C., and Jiang, J., 2014. Evaluating total inorganic nitrogen in coastal waters through fusion of multi-temporal RADARSAT-2 and optical imagery using random forest algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 33, 192–202.

Liu, H., Weng, Q., 2012. Enhancing temporal resolution of satellite imagery for public health studies: A case study of west nile virus outbreak in Los Angeles in 2007. *Remote Sensing of Environment*, 117, 57–71.

Loncan, L., De Almeida, L. B., Bioucas-Dias, J. M., Briottet, X., Chanussot, J., Dobigeon, N., Fabre, S. et al., 2015. Hyperspectral pansharpening: A review. *IEEE Geoscience and Remote Sensing Magazine*, 3, 27–46.

Lu, L., Xie, W., Zhang, J., Huang, G., Li, Q., and Zhao, Z., 2015. Woodland extraction from high-resolution CASMSAR data based on Dempster-Shafer evidence theory fusion. *Remote Sensing*, 7, 4068–4091.

Mascarenhas, N. D. A., Banon, G. J. F., and Candeias, A. L. B., 1996. Multispectral image data fusion under a Bayesian approach. *International Journal of Remote Sensing*, 17, 1457–1471.

Masi, G., Cozzolino, D., Verdoliva, L., and Scarpa, G., 2016. Pansharpening by convolutional neural networks. *Remote Sensing*, 8, 594.

Melanie, M., 1998. *An Introduction to Genetic Algorithms*. The MIT Press, Cambridge, MA, USA.

Melgani, F., 2004. Classification of multitemporal remote-sensing images by a fuzzy fusion of spectral and spatio-temporal contextual information. *International Journal of Pattern Recognition and Artificial Intelligence*, 18, 143–156.

Meng, Q., Borders, B., and Madden, M., 2010. High-resolution satellite image fusion using regression kriging. *International Journal of Remote Sensing*, 31, 1857–1876.

Mitchell, H. B., 2007. *Multi-Sensor Data Fusion: An Introduction*. Springer, Berlin, Heidelberg, Germany.

Mohammadzadeh, A. and Zoej, J. V., 2006. Road extraction based on fuzzy logic and mathematical morphology from pan-sharpened ikonos images. *Photogrammetric Record*, 21, 44–60.

Moosavi, V., Talebi, A., Mokhtari, M. H., Shamsi, S. R. F., and Niazi, Y., 2015. A wavelet-artificial intelligence fusion approach (WAIFA) for blending Landsat and MODIS surface temperature. *Remote Sensing of Environment*, 169, 243–254.

Mumtaz, A., Majid, A., and Mumtaz, A., 2008. Genetic algorithms and its application to image fusion. In: *2008 IEEE the 4th International Conference on Emerging Technologies*, 6–10.

Nguyen, H., Cressie, N., and Braverman, A., 2012. Spatial statistical data fusion for remote sensing applications. *Journal of the American Statistical Association*, 107, 1004–1018.

Nguyen, H., Franke, K., and Petrovic, S., 2009. Optimizing a class of feature selection measures. In: *NIPS 2009 Workshop on Discrete Optimization in Machine Learning: Submodularity, Sparsity & Polyhedra (DISCML)*, Vancouver, Canada.

Nishii, R., 2003. A markov random field-based approach to decision-level fusion for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 2316–2319.

Nishii, R., Kusanobu, S., and Tanaka, S., 1996. Enhancement of low spatial resolution image based on high resolution bands. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 1151–1158.

Palsson, F., Sveinsson, J. R., Ulfarsson, M. O., 2017. Multispectral and hyperspectral image fusion using a 3-D-convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 14, 639–643.

Pardo-Igúzquiza, E., Chica-Olmo, M., and Atkinson, P. M., 2006. Downscaling cokriging for image sharpening. *Remote Sensing of Environment*, 102, 86–98.

Park, J. H. and Kang, M. G., 2004. Spatially adaptive multi-resolution multispectral image fusion. *International Journal of Remote Sensing*, 25, 5491–5508.

Peng, X. and Dang, A., 2010. Hybrid genetic algorithm (GA)-based neural network for multispectral image fusion. In: *2010 IEEE International Geoscience and Remote Sensing Symposium*, 496–498.

Pohl, C. and Van Genderen, J. L., 1998. Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19, 823–854.

Price, J. C., 1987. Combining panchromatic and multispectral imagery from dual resolution satellite instruments. *Remote Sensing of Environment*, 21, 119–128.

Price, J. C., 1999. Combining multispectral data of differing spatial resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1199–1203.

Puttaswamy, S. J., Nguyen, H. M., Braverman, A., Hu, X., and Liu, Y., 2013. Statistical data fusion of multi-sensor AOD over the continental United States. *Geocarto International*, 29, 48–64.

Raghavendra, R., Dorizzi, B., Rao, A., and Hemantha Kumar, G., 2011. Particle swarm optimization based fusion of near infrared and visible images for improved face verification. *Pattern Recognition*, 44, 401–411.

Rahmani, S., Strait, M., Merkurjev, D., Moeller, M., and Wittman, T., 2010. An adaptive IHS pan-sharpening method. *IEEE Geoscience and Remote Sensing Letters*, 7, 746–750.

Ranchin, T. and Wald, L., 2000. Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66, 49–61.

Rawlins, G. J. E. *(editor)*, 1991. *Foundations of Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA.

Saeedi, J. and Faez, K., 2011. A new pan-sharpening method using multiobjective particle swarm optimization and the shiftable contourlet transform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 365–381.

Sales, M. H. R., Souza, C. M., and Kyriakidis, P. C., 2013. Fusion of MODIS images using kriging with external drift. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 2250–2259.

Schowengerdt, R. A., 2006. *Remote Sensing: Models and Methods for Image Processing*. Academic Press, Cambridge, Massachusetts, USA; 3 edition (September 11, 2006)

Siddiqui, A. B., Arfan Jaffar, M., Hussain, A., and Mirza, A. M., 2011. Block-based pixel level multi-focus image fusion using particle swarm optimization. *International Journal of Innovative Computing, Information and Control*, 7, 3583–3596.

Settle, J. J. and Drake, N. A., 1993. Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14, 1159–1177.

Shafer, G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Victoria.

Shah, V. P., Younan, N. H., and King, R. L., 2008. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1323–1335.

Shettigara, V. K., 1992. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. *Photogrammetric Engineering and Remote Sensing*, 58, 561–567.

Shi, Y., Zhou, X., Yang, X., Shi, L., and Ma, S., 2015. Merging satellite ocean color data with Bayesian maximum entropy method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3294–3304.

Solaiman, B., Pierce, L. E., and Ulaby, F. T., 1999. Multisensor data fusion using fuzzy concepts: Application to land-cover classification using ERS-1/JERS-1 SAR composites. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1316–1326.

Solberg, A. H. S., Jain, A. K., and Taxt, T., 1994. Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 32, 768–778.

Soliman, O. S., Mahmoud, A. S., and Hassan, S. M., 2012. Remote sensing satellite images classification using support vector machine and particle swarm optimization. In: *2012 IEEE Third International Conference on Innovations in Bio-Inspired Computing and Applications*, 280–285.

Song, H. and Huang, B., 2013. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 1883–1896.

Souza, C. M., Firestone, L., Silva, L. M., and Roberts, D., 2003. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment*, 87, 494–506.

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Al-Shrafany, D., and Islam, T., 2013. Data fusion techniques for improving soil moisture deficit using SMOS satellite and WRF-NOAH land surface model. *Water Resources Management*, 27, 5069–5087.

Sun, J., Zhu, H., Xu, Z., and Han, C., 2013. Poisson image fusion based on Markov random field fusion model. *Information Fusion*, 14, 241–254.

Tang, Q., Bo, Y., and Zhu, Y., 2016. Spatiotemporal fusion of multiple-satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method. *Journal of Geophysical Research: Atmospheres*, 121, 4034–4048.

Tewes, A., Thonfeld, F., Schmidt, M., Oomen, R. J., Zhu, X., Dubovyk, O., Menz, G., and Schellberg, J., 2015. Using RapidEye and MODIS data fusion to monitor vegetation dynamics in semi-arid rangelands in South Africa. *Remote Sensing*, 7, 6510–6534.

Thomas, C., Ranchin, T., Wald, L., and Chanussot, J., 2008. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1301–1312.

Toet, A., 1989. A morphological pyramidal image decomposition. *Pattern Recognition, Pattern Recognition Letters*, 9, 255–261.

Tsai, V. J. D., 2003. Frequency-based fusion of multiresolution images. In: *2003 IEEE International Geoscience and Remote Sensing Symposium*, 3665–3667.

Tu, T.-M., Huang, P. S., Hung, C.-L., and Chang, C.-P., 2004. A fast intensity–hue–saturation fusion technique with spectral adjustment for IKONOS imagery. *IEEE Geoscience and Remote Sensing Letters*, 1, 309–312.

Tu, T.-M., Su, S.-C., Shyu, H.-C., and Huang, P. S., 2001. A new look at IHS-like image fusion methods. *Information Fusion*, 2, 177–186.

Tupin, F., Bloch, I., and Maitre, H., 1999. A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1327–1343.

Valdés Hernández, M. del C., and Inamura, M., 2000. Spatial resolution improvement of remotely sensed images by a fully interconnected neural network approach. *IEEE Transactions on Geoscience and Remote Sensing*, 38, 2426–2430.

Vandenberghe, L. and Boyd, S., 1996. Semidefinite Programming. *Society for Industrial and Applied Mathematics*, 38, 49–95.

van der Wal, D. and Herman, P. M. J., 2007. Regression-based synergy of optical, shortwave infrared and microwave remote sensing for monitoring the grain-size of intertidal sediments. *Remote Sensing of Environment*, 111, 89–106.

Vapnick, V. N., 1998. *Statistical Learning Theory.* Wiley, New York.

Vijayaraj, V., O'Hara, C. G., and Younan, N. H., 2004. Quality analysis of pansharpened images. In: *2004 IEEE International Geoscience and Remote Sensing Symposium*, 1–4.

Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G. A., Restaino, R., and Wald, L., 2015. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 53, 2565–2586.

Wald, L., Ranchin, T., and Mangolini, M., 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric Engineering and Remote Sensing*, 63, 691–699.

Walker, J. J., de Beurs, K. M., Wynne, R. H., and Gao, F., 2012. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment*, 117, 381–393.

Wang, H., Fan, T., Shi, A., Huang, F., and Wang, H., 2010. Fuzzy integral based information fusion for water quality monitoring using remote sensing data. *International Journal of Communications, Network and System Sciences*, 3, 737–744.

Wang, Q., Shi, W., and Atkinson, P. M., 2016. Area-to-point regression kriging for pan-sharpening. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 151–165.

Wang, Q., Shi, W., Atkinson, P. M., and Wei, Q., 2017a. Approximate area-to-point regression kriging for fast hyperspectral image sharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10, 286–295.

Wang, Q., Shi, W., Atkinson, P. M., and Zhao, Y., 2015a. Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, 166, 191–204.

Wang, W., Jiao, L., and Yang, S., 2014. Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model. *Information Fusion*, 20, 73–87.

Wang, W., Jiao, L., and Yang, S., 2015b. Novel adaptive component-substitution-based pan-sharpening using particle swarm optimization. *IEEE Geoscience and Remote Sensing Letters*, 12, 781–785.

Wang, Q., Zhang, Y., Onojeghuo, A. O., Zhu, X., and Atkinson, P. M., 2017b. Enhancing spatio-temporal fusion of MODIS and Landsat data by incorporating 250 m MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10, 4116–4123.

Wang, Z., Ziou, D., Armenakis, C., Li, D., and Li, Q., 2005. A comparative analysis of image fusion methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 1391–1402.

Waske, B. and Benediktsson, J. A., 2007. Fusion of support vector machines for classification of multisensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 3858–3866.

Wei, J., Wang, L., Liu, P., and Song, W., 2016. Spatiotemporal fusion of remote sensing images with structural sparsity and semi-coupled dictionary learning. *Remote Sensing*, 9, 21.

Welch, R. and Ehlers, M., 1987. Merging multiresolution SPOT HRV and Landsat TM data. *Photogrammetric Engineering and Remote Sensing*, 53, 301–303.

Weng, Q., Fu, P., and Gao, F., 2014. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment*, 145, 55–67.

Wong, F. H. and Orth, R., 1980. Registration of SEASAT/LANDSAT composite images to UTM coordinates, In: Proceedings of the Sixth Canadian Syinposium on Remote Sensing, 161–164.

Wu, M., Niu, Z., Wang, C., Wu, C., and Wang, L., 2012. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model. *Journal of Applied Remote Sensing*, 6, 63507.

Xu, M., Chen, H., and Varshney, P. K., 2011. An image fusion approach based on Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 49, 5116–5127.

Xu, H., Guang, J., Xue, Y., de Leeuw, G., Che, Y. H., Guo, J., He, X. W., and Wang, T. K., 2015a. A consistent aerosol optical depth (AOD) dataset over mainland China by integration of several AOD products. *Atmospheric Environment*, 114, 48–56.

Xu, Y., Huang, B., Xu, Y., Cao, K., Guo, C., and Meng, D., 2015b. Spatial and temporal image fusion via regularized spatial unmixing. *IEEE Geoscience and Remote Sensing Letters*, 12, 1362–1366.

Yang, X.-S., Deb, S., and Fong, S., 2011. Accelerated particle swarm optimization and support vector machine for business optimization and applications. In: *Networked Digital Technologies: Third International Conference, NDT 2011*, 53–66. Macau, China, July 11–13, 2011.

Yang, H., Du, Q., and Chen, G., 2012a. Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 544–554.

Yang, B. and Li, S., 2010. Multifocus image fusion and restoration with sparse representation. *IEEE Transactions on Instrumentation and Measurement*, 59, 884–892.

Yang, G., Pu, R., Huang, W., Wang, J., and Zhao, C., 2010. A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 2170–2178.

Yang, S., Wang, M., and Jiao, L., 2012b. Fusion of multispectral and panchromatic images based on support value transform and adaptive principal component analysis. *Information Fusion*, 13, 177–184.

Yang, S., Zeng, L., Jiao, L., and Xiao, J., 2007. Fusion of multispectral and panchromatic images using improved GIHS and PCA mergers based on contourlet. In: *Proceedings of the SPIE*, Volume 6787, Multispectral Image Processing.

Yokoya, N., Yairi, T., and Iwasaki, A., 2012. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 528–537.

Zadeh, L. A., 1965. Fuzzy sets. *Information and Control*, 8, 338–353.

Zhan, Z. H., Zhang, J., Li, Y., and Chung, H. S.-H., 2009. Adaptive particle swarm optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39, 1362–1381.

Zhan, Z.-H., Zhang, J., Li, Y., and Shi, Y.-H., 2011. Orthogonal learning particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, 15, 832–847.

Zhang, J., 2010. Multi-source remote sensing data fusion: Status and trends. *International Journal of Image and Data Fusion*, 1, 5–24.

Zhang, Y., Atkinson, P. M., Ling, F., Wang, Q., Li, X., Shi, L., and Du, Y., 2017. Spectral–spatial adaptive area-to-point regression Kriging for MODIS image downscaling. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10, 1883–1896.

Zhang, Q. and Guo, B. L., 2009. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Processing*, 89, 1334–1346.

Zhang, H. and Huang, B., 2015. A new look at image fusion methods from a Bayesian perspective. *Remote Sensing*, 7, 6828–6861.

Zhang, W., Li, A., Jin, H., Bian, J., Zhang, Z., Lei, G., Qin, Z., and Huang, C., 2013. An enhanced spatial and temporal data fusion model for fusing Landsat and MODIS surface reflectance to generate high temporal Landsat-like data. *Remote Sensing*, 5, 5346–5368.

Zhang, F., Zhu, X., and Liu, D., 2014. Blending MODIS and Landsat images for urban flood mapping. *International Journal of Remote Sensing*, 35, 3237–3253.

Zhong, J., Yang, B., Huang, G., Zhong, F., and Chen, Z., 2016. Remote sensing image fusion with convolutional neural network. *Sensing and Imaging*, 17, 10.

Zhou, J., Civco, D. L., and Silander, J. A., 1998. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, 19, 743–757.

Zhou, J., Liu, L., Guo, J., and Sun, L., 2013. Multisensor data fusion for water quality evaluation using Dempster-Shafer evidence theory. *International Journal Of Distributed Sensor Networks*, 2013, 1–6.

Zhu, X., Chen, J., Gao, F., Chen, X., and Masek, J. G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114, 2610–2623.

Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., and Lefsky, M. A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment*, 172, 165–177.

Zhukov, B., Oertel, D., Lanzl, F., and Reinhäckel, G., 1999. Unmixing-based multisensor multiresolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1212–1226.

Zurita-Milla, R., Clevers, J. G. P. W., and Schaepman, M. E., 2008. Unmixing-based landsat TM and MERIS FR data fusion. *IEEE Geoscience and Remote Sensing Letters*, 5, 453–457.

Zurita-Milla, R., Clevers, J. G. P. W., Van Gijsel, J. A. E., and Schaepman, M. E., 2011. Using MERIS fused images for land-cover mapping and vegetation status assessment in heterogeneous landscapes. *International Journal of Remote Sensing*, 32, 973–991.

Zurita-Milla, R., Kaiser, G., Clevers, J. G. P. W., Schneider, W., and Schaepman, M. E., 2009. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sensing of Environment*, 113, 1874–1885.