PRISMA Spatial Resolution Enhancement by Fusion With Sentinel-2 Data

Nicola Acito, *Member, IEEE*, Marco Diani, *Member, IEEE*, and Giovanni Corsini, *Member, IEEE*

***Abstract*—This article deals with the problem of improving the spatial resolution of hyperspectral (HS) data from the PRecursore IperSpettrale della Missione Applicativa (PRISMA) mission. For this purpose, higher spatial resolution data from the Sentinel-2 (S2) mission are exploited. Particularly, 10 S2 bands at 10 and 20 m spatial resolution are used to accomplish the PRISMA super-resolution (SR) task. The article presents a new end-to-end procedure, called PRISMA-SR, that starting from the S2 data and the low-resolution PRISMA image, provides a super-resolved image with a spatial resolution of 10 m and the same spectral resolution as the PRISMA HS sensor. The first step of the PRISMA-SR procedure consists in fusing S2 data at different spatial resolutions to obtain a synthetic MS image with 10 m spatial resolution and 10 spectral bands. Then, an unsupervised procedure is applied to coregister the fused S2 image and the PRISMA image. Finally, the two images at different spatial resolutions are properly combined in order to obtain the super-resolved HS image. Solutions for each step of the PRISMA-SR processing chain are proposed and discussed. Simulated data are used to show the effectiveness of the PRISMA-SR scheme and to investigate the impact on its performance of each step of the processing chain. Real S2 and PRISMA images are finally considered to provide an example of the application of the PRISMA-SR.**

***Index Terms*—Hyperspectral (HS) data processing, hyperspectral (HS)-multispectral (MS) data fusion, satellite missions.**

I. Introduction

HYPERSPECTRAL (HS) sensors offer the opportunity of analyzing the chemical and physical composition of the remotely sensed scene, thanks to their ability of measuring the spectrum of the observed pixels in a large number of contiguous and narrow spectral channels [1]. In particular, spaceborne sensors allow the exploitation of HS technology for large-scale monitoring of the earth [2]. The set of Earth Observation (EO) applications enabled by spaceborne HSs includes [3]–[6] detailed environmental monitoring, forest analysis, precision agriculture, inland and coastal water monitoring, raw material exploration and mining, and soil degradation and analysis.

Manuscript received July 5, 2021; revised September 17, 2021 and October 22, 2021; accepted November 29, 2021. Date of publication December 2, 2021; date of current version December 20, 2021. This work was supported by the PRISMA products, © of the Italian Space Agency, delivered under an ASI License to use. *(Corresponding author: Nicola Acito.)*

Nicola Acito is with the Department of Information Engineering, University of Pisa, 56122 Pisa, Italy, and also with the Consorzio Nazionale Interuniversitario per le Telecomunicazioni, 43124 Parma, Italy (e-mail: nicola.acito@unipi.it).

Marco Diani is with the Accademia Navale, 57127 Livorno, Italy (e-mail: m.diani@iet.unipi.it).

Giovanni Corsini is with the Department of Information Engineering, University of Pisa, 56122 Pisa, Italy (e-mail: giovanni.corsini@unipi.it).

Digital Object Identifier 10.1109/JSTARS.2021.3132135

In this context, the EO mission of the Italian Space Agency (ASI) named PRISMA (PRecursore IperSpettrale della Missione Applicativa, [7]–[10]) offers a great opportunity to improve the knowledge about the scientific and commercial applications of spaceborne HS data. The mission has been operational since March 2019 and data for the community are now available according to the ASI data policy [11]. PRISMA payload includes a pushbroom HS camera operating in the visible, near infrared (VNIR) and shortwave infrared (SWIR) spectral ranges. Specifically, the HS imaging sensor covers the portion of the electromagnetic spectrum ranging from 400 to 2500 nm with 10 nm spectral sampling through two partially overlapped spectrometers [10]. The total number of distinct bands is 230. The satellite has a sun-synchronous circular orbit with a repeat cycle of 29 days and relook capability of 7 days with roll maneuver. Similar to other HS satellite missions the PRISMA HS sensor has a limited spatial resolution because of the limited amount of incident energy. Specifically, its ground sampling distance (GSD) is of 30 m. The low spatial resolution can reduce the range of the potential applications of the acquired data. Such applications include high-spatial-resolution ecosystem monitoring, or high-spatial-resolution mapping of minerals, urban surface materials, plant species, and many others.

A natural solution to enhance the spatial resolution of a satellite HS image is to fuse it with a higher resolution multispectral (MS) image of the same scene, possibly acquired under the same conditions. This approach is called HS and MS image fusion (HS–MS fusion), or HS super-resolution (SR). HS–MS fusion has attracted the interest of the scientific community in recent years as witnessed by numerous published papers focused on this topic [12]–[33].

However, PRISMA mission payload does not include a high-spatial resolution MS sensor. Therefore, the only way to apply the HS–MS fusion paradigm to PRISMA data is to exploit the high-resolution MS image provided by other satellite missions. In this context, a great opportunity is offered by the Sentinel-2 (S2) program.

S2 is a wide-swath and fine spatial resolution satellite imaging mission of the European Space Agency (ESA) developed in the framework of the European Union Copernicus program [24], [35]. Sentiel-2 consists of two satellite instruments, Sentinel 2A (S2A) and Sentinel 2B (S2B), which are identical and have the same sun-synchronous, quasi-circular, near-polar, and low-earth orbit. Each sensor has a 10-day repeat cycle, and together they provide globally a median average satellite revisit of 3.7 days [35]. S2A and S2B acquire data with GSD of 10, 20, and 60

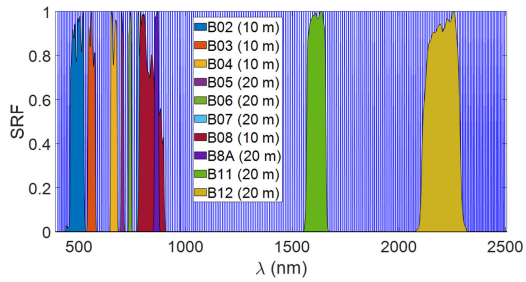


Fig. 1. Center wavelengths of PRISMA bands (blue bars) and SRFs of the Sentinel 2 bands with 10 and 20 m spatial resolution.

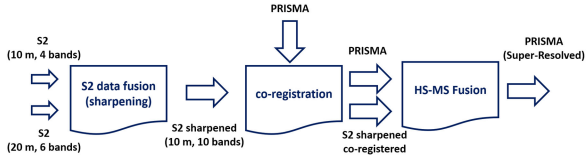


Fig. 2. PRISMA super-resolution scheme based on Sentinel 2 fusion.

m in the VNIR and SWIR spectral ranges [24]. Specifically, the 10 bands with spatial resolution of 10 m (B2–B4 and B8) and 20 m (B5–B7, B8a, and B11–B12) cover a spectral range which is overlapped with that of the PRISMA sensor and can be exploited to enhance the spatial resolution of PRISMA data according to the HS–MS fusion framework. For completeness, in Fig. 1, we show the overlap among the nominal spectral response functions (SRFs) of the 10 and 20 m S2 bands [36] and those of the PRISMA sensor.

In this article, we focus on the problem of fusing PRISMA and S2 reflectance images which, to the best of our knowledge, has not been yet investigated in the open literature. Here, fusion aims at obtaining a synthetic 10-m spatial resolution HS image with the same spectral resolution as PRISMA. Though the general idea of fusing HS and MS data has been already investigated in the literature [12]–[33], here we propose a novel end-to-end paradigm for PRISMA and S2 data, discussing all the steps necessary to get the fused image and their impact on the fusion output by means of simulated data.

As is shown in Fig. 2, the proposed PRISMA-SR strategy consists of three main steps. The first step (named S2 data fusion) applies a fusion algorithm to the 10 and 20 m S2 images to obtain a 10 band MS image with the highest spatial resolution which will be used to bring the spatial resolution of the PRISMA image to 10 m. The second step (coregistration) applies a registration procedure to the MS image resulting from the previous step and the PRISMA image. This step is necessary even in the case of geolocated data to compensate the inaccuracy of pixel-level coordinates due to errors in the georeferencing process. The third step (HS–MS fusion) finally applies an HS–MS fusion scheme to obtain a *super-resolved* PRISMA image. Of course, the PRISMA and S2 images must be acquired on the same scene and with the minimum time difference in order to limit the potential seasonal changes in the monitored surface.

The main goal of this article is to test the effectiveness of the general PRISMA-SR scheme of Fig. 2, proposing simple solutions for each step. As to the S2 data fusion step, among the several methods proposed since 2016 [37] to fuse the 10 and the 20 m S2 bands [38]–[41], we have chosen a multiresolution analysis (MRA) based algorithm (detailed in Section III) for its computational efficiency and its good fusion performance. Image coregistration is performed assuming an affine transformation between the images and exploiting control points (CPs) detected in an unsupervised manner by the speeded-up robust feature (SURF, [42]) algorithm. HS–MS fusion is surely the most critical step of the PRISMA-SR scheme. Several algorithms (summarized in Section II) have been proposed in the past years. Most of them assume perfect coregistration of the two images, a hypothesis that is difficult to satisfy in practice. In fact, regardless of the adopted coregistration algorithm, residual registration errors still remain that can deteriorate the performance of the HS–MS fusion procedure. For this reason, we propose two modified versions of the well-known *smoothing filtered-based intensity modulation* (SFIM) algorithm [14], [43] specifically designed to be robust to residual registration errors. An extensive analysis on synthetic datasets is carried out to investigate the performance of the proposed PRISMA-SR scheme and the effectiveness of each step of its processing chain. Furthermore, real PRISMA and S2 data are analyzed to provide an example of the application of the PRISMA-SR procedure in a practical case.

The main contributions of this article are summarized as follows:

1) An end-to-end procedure to fuse S2 and PRISMA data. The procedure starts from the fusion of S2 data at different spatial resolution to obtain synthetic S2 data with 10 bands and 10 m spatial resolution. A coregistration algorithm is then applied to the S2 synthetic image and the PRISMA image. Finally, an HS–MS fusion algorithm is applied to obtain the super-resolved PRISMA image.
2) The idea of combining S2 data at two different spatial resolutions for PRISMA data spatial enhancement and the analysis of its effectiveness by means of simulated data.
3) Two modified versions of the well-known SFIM fusion algorithm that mitigate the effects of residual errors in the image registration process.

The rest of the article is organized as follows. Section II briefly reviews the methods presented in the literature for HS–MS fusion, which represents the most critical step of the proposed PRISMA-SR scheme. Section III details the solution proposed for each step of PRISMA-SR scheme paying particular attention to the HS–MS fusion algorithm. Section IV illustrates the simulation-based performance evaluation procedure and summarizes the results obtained. Finally, Section V shows the results obtained on a real PRISMA and S2 image pair.

II. HS–MS Fusion Algorithms

In recent years, several HS–MS fusion approaches have been proposed. They can be grouped into three categories: pansharpening adaptation methods, subspace-based approaches, and machine learning algorithms.

The first category includes methods obtained by adapting the techniques developed to fuse a high-resolution panchromatic (PAN) image with a lower resolution MS image (pansharpening [12]) to the HS–MS fusion problem. In this context, Selva *et al.* [13] proposed a framework [called hypersharpening (HP)] that adapts pansharpening algorithms based on the MRA [12] to HS–MS fusion. Such framework consists in synthesizing a high-resolution image for each HS band as a linear combination of the MS bands. The linear combination coefficients are derived through linear regression. The authors also proposed an alternative approach (called band-selected) that allows the adaptation of the pansharpening methods based on component substitution (CS) to the HS–MS fusion. In this approach, the most correlated MS band is selected for each HS band to reduce the HS–MS fusion problem to distinct pansharpening problems that can be attacked by one of the available pansharpening algorithms included the CS based ones. In [14], the band-selected scheme was used to adapt the Gram–Schmidt Adaptive [14] algorithm to the HS–MS fusion. HP paradigm was also proposed [14] to adapt the SFIM method to the HS–MS fusion problem. The resulting algorithm was tested on several scenarios in [14].

The subspace-based category includes the unmixing [16]–[20] and the Bayesian-based [21]–[24] methods. The former associate HS–MS fusion to a linear spectral unmixing problem, modeling the data as a mixture of a relatively small number of "pure" spectral signatures (endmembers), corresponding to the materials present in the scene. HS–MS fusion is performed by obtaining endmember information from the HS image and high-resolution abundance map (representing the coefficients of the mixture) from the MS image. The extraction of endmembers and abundances accounts for both the SRFs and the Modulation Transfer Functions (MTFs) of the considered HS and MS sensors. The final result is obtained as the product of the endmembers and the high-resolution abundance map. In [16], the coupled subspace nonnegative matrix factorization algorithm was proposed that incorporates the relative SRF and MTF of the HS and MS sensors in an iterative scheme based on nonnegative matrix factorization aimed at estimating both the endmember and the abundance map. Based on the assumption that the number of endmembers in each pixel is small compared to the total number of endmembers in the analyzed data, sparse regularization was exploited in [17] and [18] to perform unmixing. In [19], the problem of estimating the endmembers and the abundance map from HS and MS images was addressed by including the nonnegative constraint on both the endmembers and the abundance map and the sum-to-one constraint on the abundance map in order to respect physical constraints. Local processing was exploited in [20] to mitigate the ill-posedness of the fusion problem.

Bayesian-based approaches incorporate the idea underlying mixture-based methods into a Bayesian estimation framework by introducing a regularization term in the objective function based on some priors in order to address the ill-posedness of the fusion problem. In general, Gaussian or sparsity promoted Gaussian priors are adopted [21], [22] and singular value decomposition or vertex component analysis [44] performed on the HS input image are exploited to address the data subspace. A computationally efficient Bayesian-based method was proposed in [23] which approaches the fusion problem by exploiting the closed-form solution of a Sylvester equation. Regularization based on vector total variation is incorporated in the Bayesian formulation of the HS–MS fusion problem. The problem is thus recasted as a convex optimization problem and approached by the alternating direction method of multipliers (ADMM). The resulting algorithm, named HySure, includes the estimation of the relative SRFs and MTFs of the two images and has provided excellent results in a variety of multisource HS–MS fusion problems [14].

A very interesting comparative review concerning the categories of algorithms summarized above can be found in [14].

Machine learning and deep learning are emerging approaches that have received more and more attention in many image processing applications. In the context of image fusion, in the last years, several convolutional neural network (CNN) based algorithms have been proposed to approach the pansharpening problem [26]–[28]. Inspired by the architecture proposed in [25] for single image SR, in [26], a CNN-based algorithm has been developed for the fusion of PAN and MS images by augmenting the inputs of the CNN in [25] including a number of features extracted from the images to be fused. More recently [27], a generative adversarial network has been proposed for the fusion of PAN and MS images. The network exploits two discriminators to force the spectral and spatial information of the generated super-resolved image to be consistent with the MS image and the PAN image, respectively. In [28], a new residual learning network for pansharpening has been proposed. It consists of two networks: the gradient transformation network and the pansharpening network. The first one learns the gradient mapping between the MS image bands and the PAN image in both the spatial directions. The second network adopts residual learning to perform the fusion of the MS and the PAN images. The spatial details of the super-resolved MS image are derived not only on the basis of the PAN image but also on the gradient transformation relationship established by the first network.

In recent years, CNN-based algorithms have also been proposed for the HS–MS fusion problem. Dian *et al.* [29] proposed a CNN-based fusion scheme that exploits both the potential of residual learning and of the Bayesian approach. The method first initializes a high spatial resolution (HR) HS image from the model based fusion framework solving a Sylvester equation. Then, it derives an intermediate HR–HS image exploiting a CNN based on residual learning. This intermediate HR–HS image is returned to the model-based fusion framework to obtain the final super-resolved HS image. Xie *et al.* [30] used a deep network to capture deep residual information from high-frequency domain and used it as extra constraint for reconstructing the super-resolved HS image. Han *et al.* [31] proposed a multiscale deep CNN to gradually improve the resolution of the original HS image. The proposed fusion scheme integrates a multilevel cost function to alleviate the problem of gradient vanishing in the training phase. In [32], a fusion scheme that connects the approach based on deep learning and the one based on Bayesian optimization is proposed. The Bayesian optimization problem is reformulated as three suboptimization problems modeled by

TABLE I
LIST OF THE PRINCIPAL ACRONYMS

Acronym	Meaning
ASI	Italian Space Agency
CBD	Context Based Decision
CP	Control Point
ESA	European Space Agency
FC	False Color
FWHM	Full Width Half Maximum
GLP	Generalized Laplacian Pyramid
GSD	Ground Sampling Distance
HP	Hyper-sharPening
HS	HyperSpectral
HS-SR	Hyperspectral Super-Resolution
IP	Interesting Point
MRA	Multi Resolution Analysis
MS	MultiSpectral
MTF	Modulation Transfer Function
OLS	Ordinary Least Squares
PRISMA	Precursore IperSpettrale della Missione Applicativa
PRISMA-SR	PRISMA-Super Resolution
PSNR	Peak Signal to Noise Ratioi
RFSIM-HP-	Robust SFIM with Hyper-sharpPening
RGB	Red Green Blue
RRMSE	Relative Root Mean Squared Error
RSFIM	Robust SFIM
S2	Sentinel-2
SAM	Spectral Angle Mapper
SFIM	Smoothing Filtered-based Intensity Modulation
SFIM-HP	SFIM with Hyper-sharPening
SNR	Signal to Noise Ratio
SR	Super Resolution
SRF	Spectral Response Function
SURF	Speeded Up Robust Feature
SWIR	Short Wave InfraRed
VNIR	Visible and Near InfraRed
PAN	Panchromatic

TABLE II
NOTATION

Symbol	Meaning
$\boldsymbol{X}'_{10}$	Nxl_{10} Sentinel 2 image at 10 m (bands: B2-B4, B8)
$\boldsymbol{X}'_{20}$	N_{20}xl_{20} Sentinel 2 image at 20 m (bands: B5-B7, B8a, B11-B12)
$\boldsymbol{X}'$	Nxl Sentinel 2 sharpened image at 10 m (bands: B2-B4, B5-B7, B8, B8a, B11-B12)
$\boldsymbol{Y}$	nxL PRISMA image at 30 m (VNIR-SWIR)
$\boldsymbol{X}$	Nxl Sentinel sharpened image at 10 m co-registered to the PRISMA image
$\boldsymbol{Z}$	NxL PRISMA super-resolved image
$\mathbf{S}$	lxL spectral response matrix, each row holds the spectral response function of the corresponding S2 band resampled at the spectral resolution of the PRISMA sensor.
$h_G(i,j;c)$	Gaussian filter impulse response whose amplitude response value at the Nyquist frequency is equal to c
c	value at the Nyquist frequency of the sensor Modulation Transfer Function
$\sim$	Operator indicating the down-scaling at the MS spectral resolution
$(\ )^{\uparrow}$	Operator indicating the spatial up-sampling
$(\ )^{\downarrow}$	Operator indicating the spatial down-sampling
$(\ )_{LP}$	Operator indicating the Low Pass filtering

neural network architectures. In particular, the powerful representation ability of a deep neural network is exploited to replace the priors generally used in the Bayesian approach with a more complex function of the data.

Although the CNN-based approaches have shown promising results, two main aspects still limit their use in practical remote sensing applications. CNN-based approaches require a large number of images for the training phase, which are often not available. The generalization ability of the deep CNN is also a critical issue that needs to be further investigated. CNN trained on one kind of data may not work well on other data types.

A comprehensive review of the recent advances on CNN-based HS–MS fusion methods can be found in [33].

III. PRISMA-SR SCHEME

As shown in Fig. 2, the proposed PRISMA-SR scheme consists of three main processing steps: S2 data fusion, coregistration, and HS–MS fusion. In this section, we describe the algorithms adopted in each block of the PRISMA-SR processing chain. Several symbols will be used to denote the data involved in each algorithm. A quick summary of the main symbols used in the text is listed in Table II.

We will adopt the lexicographic notation for the MS and HS images. The generic image $\boldsymbol{\Psi}$ will be represented as an $N \times B$ matrix ($\boldsymbol{\Psi} \in \mathcal{R}^{N\times B}$), with N denoting the number of image pixels and B the number of bands. However, whenever needed, we will use the notation $\boldsymbol{\Psi}(i,j) \in \mathcal{R}^{B\times 1}$ to explicitly indicate the spatial coordinates i and j of the spectral pixel. Furthermore, we will use the symbol $\Psi_b \in \mathcal{R}^{N\times 1}$ to denote the N pixels image in the bth spectral band.

A. S2 Data Fusion

As to the S2 data fusion, we consider the HP adaptation of the MTF–Gaussian Laplacian Pyramid (GLP)–context-based decision (CBD) algorithm proposed in [45].

MTF–GLP–CBD is a pansharpening algorithm belonging to the MRA class here chosen because of its computational efficiency and of the good fusion performance experienced on real S2 data [37]. In [37], to extend the algorithm to the S2 case, the HP paradigm was adopted and the best performance was obtained by the synthesized band scheme. Let us denote as $\boldsymbol{X}'_{10} \in \mathcal{R}^{N\times l_{10}}$ the S2 image with a spatial resolution of 10 m, and with $\boldsymbol{X}'_{20} \in \mathcal{R}^{N_{20}\times l_{20}}$ the S2 image with a spatial resolution of 20 m. N and N_{20} ($N = 4N_{20}$) are the number of pixels of $\boldsymbol{X}'_{10}$ and $\boldsymbol{X}'_{20}$, respectively. While $l_{10} = 4$ and $l_{20} = 6$ are the number of bands of $\boldsymbol{X}'_{10}$ and $\boldsymbol{X}'_{20}$, respectively. The synthesized band HP paradigm starts from a HR version of $\boldsymbol{X}'_{20}$ ($\boldsymbol{P}_S \in \mathcal{R}^{N_{10}\times l_{20}}$) obtained by linear regression from $\boldsymbol{X}'_{10}$:

$$\boldsymbol{P}_S = \boldsymbol{X}'_{10} \cdot \boldsymbol{M}_S + \boldsymbol{b}_S \tag{1}$$

where $\boldsymbol{M}_S \in \mathcal{R}^{l_{10}\times l_{20}}$ and $\boldsymbol{b}_S \in \mathcal{R}^{1\times l_{20}}$ are obtained by ordinary least squares (OLS) applied to $\boldsymbol{X}'_{20}$ and $(\boldsymbol{X}'_{10})^{\downarrow}$. The latter

being the version of $\boldsymbol{X}'_{10}$ downsampled by a factor of 2 so as to get the same spatial resolution of $\boldsymbol{X}'_{20}$. MTF–GLP–CBD is then applied to obtain the super-resolved version $\hat{\boldsymbol{X}}' \in \mathcal{R}^{N\times l_{20}}$ of $\boldsymbol{X}'_{20}$ with each band of $\boldsymbol{P}_S$ acting as the PAN image. In formulas:

$$\hat{\boldsymbol{X}}' = (\boldsymbol{X}'_{20})^{\uparrow} + [\boldsymbol{P}_S - (\boldsymbol{P}_S)_{\mathrm{LP}}] \cdot \boldsymbol{G} \tag{2}$$

where $(\boldsymbol{X}'_{20})^{\uparrow}$ is obtained by interpolating[1] $\boldsymbol{X}'_{20}$ at the fine scale, $(\boldsymbol{P}_S)_{\mathrm{LP}}$ is a low-pass version of $\boldsymbol{P}_S$ and $\boldsymbol{G} \in \mathcal{R}^{l_{20}\times l_{20}}$ is the gain matrix. $(\boldsymbol{P}_S)_{\mathrm{LP}}$ is obtained by convolving each band of $\boldsymbol{P}_S$ with a Gaussian filter approximating the sensor MTF in that band. The MTF value at the Nyquist frequency is set according to the S2 L1C data quality report [46]. The gain matrix $\boldsymbol{G}$ is set according to the CBD paradigm as a diagonal matrix, whose bth diagonal entry is the ratio between the covariance between the bth band of $(\boldsymbol{X}'_{20})^{\uparrow}$ and $(\boldsymbol{P}_S)_{\mathrm{LP}}$ and the variance of the bth band of $(\boldsymbol{P}_S)_{\mathrm{LP}}$.

Finally, the S2 sharpened image $\boldsymbol{X}' \in \mathcal{R}^{N\times l}$ with $l = 10$ is obtained by properly merging $\boldsymbol{X}'_{10}$ and $\hat{\boldsymbol{X}}'$.

B. Coregistration of S2 and PRISMA Images

Coregistration between S2 and PRISMA images is performed in an unsupervised way by using the well-established feature-based technique that exploits the SURF algorithm [47], [48]. In this subsection, we briefly summarize the basics of that technique by pointing out the modifications made to adapt it to the case of coregistration of S2 and PRISMA images.

In general, feature-based techniques start from the extraction of interest points (IPs) independently from the two images to be coregistered named test image and reference image. To each IP, a feature vector (descriptor) is then assigned by exploring the characteristics of the image in a local neighborhood. In our case, the SURF algorithm [42] is adopted to extract descriptors that are invariant to scale, rotation, and illumination. IPs from the two different images are tested to find pairwise correspondence between their descriptors. Pairs of IPs with the most similar descriptors are retained and represent the so-called CPs. Finally, the pixel coordinates of each CP in the two images are used to estimate, in accordance to a given model, the transformation that represents the geometrical distortion between the test and the reference image. Once the transformation is estimated, it is applied to resample the test image into the coordinates system of the reference image.

In our case, the reference and the test images are the PRISMA and the S2 sharpened image ($\boldsymbol{X}' \in \mathcal{R}^{N\times l}$), respectively. More precisely, the spectrally downsampled version ($\tilde{\boldsymbol{Y}} \in \mathcal{R}^{n\times l},\ N = 9n$) of the PRISMA image is adopted as reference. $\tilde{\boldsymbol{Y}}$ is obtained as

$$\tilde{\boldsymbol{Y}} = \boldsymbol{Y} \cdot \mathbf{S}^T \tag{3}$$

where $\boldsymbol{Y} \in \mathcal{R}^{n\times L}$ is the PRISMA image and $\mathbf{S}$ is the $l \times L$ relative spectral response matrix, i.e., each row of $\mathbf{S}$ is the SRF of the corresponding S2 band resampled at the spectral resolution of the PRISMA sensor. $\mathbf{S}$ is obtained by the nominal S2 SRFs provided by ESA [36].

The SURF-based procedure summarized above is independently applied to each band of the two images $\boldsymbol{X}'$ and $\tilde{\boldsymbol{Y}}$. The CPs extracted in all the bands are merged together and used to estimate the geometrical distortion assuming an affine transformation with a scale ratio of 3, according to the GSDs of the two considered images. To make the geometric model estimation procedure robust to potential outliers, the popular RANdom Sampling Consensus [49] method is adopted. The estimated parameters (rotation matrix and shift vector) are used to map $\boldsymbol{X}'$ into the coordinate system of the PRISMA image, thus obtaining the S2 sharpened and coregistered image $\boldsymbol{X} \in \mathcal{R}^{N\times l}$.

C. HS–MS Fusion: Robust SFIM Algorithms

Fusion of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is accomplished by considering the HP adaptation of the SFIM algorithm [14] called SFIM-HP, hereinafter. SFIM sharpens the low spatial resolution image by multiplying the upsampled lower resolution image by the ratio between the higher resolution image and its low-pass filtered version on a pixel-by-pixel basis. Accordingly, the PRISMA superresolved image $\boldsymbol{Z} \in \mathcal{R}^{N\times L}$ is obtained as

$$\boldsymbol{Z}(i,j) = [\boldsymbol{P}_H(i,j) \oslash (\boldsymbol{P}_H(i,j);c)_{LP}] \odot (\boldsymbol{Y})^{\uparrow}(i,j) \tag{4}$$

where i and j are the pixel coordinates, $(\boldsymbol{Y})^{\uparrow}$ is the upscaled version of $\boldsymbol{Y}$, and $\oslash$ and $\odot$ denote the elementwise division and multiplication, respectively. $\boldsymbol{P}_H \in \mathcal{R}^{N\times L}$ is the HS synthetic image obtained by linear regression from $\boldsymbol{X}$ according to the HP paradigm introduced in -Section III-A. Of course, in this case, the linear regression coefficients are derived by OLS applied to $(\boldsymbol{X})^{\downarrow}$, i.e., to the version of $\boldsymbol{X}$ downscaled to the PRISMA spatial resolution, and $\boldsymbol{Y}$. $(\boldsymbol{P}_H(i,j);c)_{\mathrm{LP}}$ is the low-pass version of $\boldsymbol{P}_H$ obtained by convolution with a Gaussian filter with value c at the Nyquist frequency chosen so as to approximate the MTF of the HS sensor:

$$(\boldsymbol{P}_H(i,j);c)_{\mathrm{LP}} = \boldsymbol{P}_H(i,j) \otimes h_G(i,j;c) \tag{5}$$

where $\otimes$ denotes the two-dimensional convolution performed on each image band and $h_G(;c)$ is the Gaussian filter impulse response. The c parameter is the value of the amplitude response of the filter at the Nyquist frequency, the same value of c is used for each band.

In general, SFIM-HP introduces relatively small spectral distortions on the super-resolved image. Conversely, it is sensitive to coregistration errors. In addition, it also requires knowledge of the parameter c which may not be available in real applications. Notice that, despite the efforts made to choose the best algorithm to geometrically align the two images, residual coregistration errors will always occur in practice and should be properly considered in the fusion strategy.

The previous considerations suggested us to develop modified versions of SFIM-HP that are more robust to residual registration errors and do not require the knowledge of c. Such robust versions of SFIM-HP will be referred to as RSFIM-HP in the rest of the article.

[1] Bicubic interpolation is considered here and in the rest of the article.

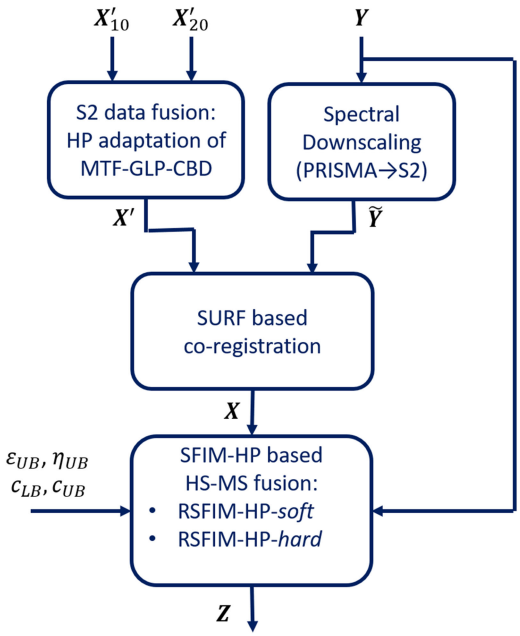


Fig. 3. Proposed PRISMA-SR scheme.

To describe the proposed RSFIM-HP algorithms, let us denote as $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ the SFIM-HP output for a given value of c and assuming that the residual registration error consists in a shift of ε and η in the two spatial directions:

$$\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c) = \left[\boldsymbol{P}_H(i,j) \oslash (\boldsymbol{P}_H(i,j);c)_{\mathrm{LP}}\right] \odot (\boldsymbol{Y})^{\uparrow}(i+\varepsilon, j+\eta). \tag{6}$$

Notice that (6) is similar to (4), but it accounts for the residual registration error in its formulation. Let us define the relative squared error between the MS HR image $\boldsymbol{X}$ and the MS image $\tilde{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c) = \mathbf{S}\cdot\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ obtained by spectral downsampling $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ at the (spectral) resolution of $\boldsymbol{X}$ as

$$E(i,j;\varepsilon,\eta,c) = \frac{\|\boldsymbol{X}(i,j) - \tilde{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)\|^2}{\|\boldsymbol{X}(i,j)\|^2}. \tag{7}$$

In the pixel of spatial coordinates (i,j), $E(i,j;\varepsilon,\eta,c)$ is expected to be low when ε, η, and c are close to their real values. Therefore, the relative squared error in (7) can be used to measure the goodness of the image $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$. The idea behind RSFIM-HP is of calculating $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ for various possible values of ε, η, and c and to combine the images $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ according to a specific optimality criterion based on the error in (7).

Specifically, we propose to derive the super-resolved HS image $(\boldsymbol{Z}(i,j))$ as a linear combination of the images $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ obtained for various values of ε, η, and c:

$$\boldsymbol{Z}(i,j) = \sum_{\substack{|\varepsilon|\le\varepsilon_{\mathrm{UB}}\\ |\eta|\le\eta_{\mathrm{UB}}\\ c_{\mathrm{LB}}<c<c_{\mathrm{UB}}}} w(i,j;\varepsilon,\eta,c)\cdot\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c) \tag{8}$$

where the weights $w(i,j;\varepsilon,\eta,c)$ for each pixel depend on the error $E(i,j;\varepsilon,\eta,c)$. In (8), $\varepsilon_{\mathrm{UB}}$ and η_{UB} (in pixel units) are the absolute maximum expected residual registration errors in the two spatial directions, while c_{LB} and c_{UB} are the lower and the upper bounds of the c parameter, respectively. The values of $\varepsilon_{\mathrm{UB}}$ and η_{UB} depend on the performance of the coregistration algorithm adopted. Typically, values in the order of few pixels are reasonable to model the residual registration error as a local shift in the two spatial directions. The sensor MTF assumes values in the range $[0,\ 1]$. Therefore, the values of c_{LB} and c_{UB} must be chosen in that range. The two bounds account for the mismatch with respect to the nominal values of the parameter c generally provided by the sensor manufacturer. Reasonable values for c_{LB} and c_{UB} are 0.2 and 0.7, respectively, which define a sufficiently large part of the parameter space by discarding the values corresponding to unlikely conditions. Low values of c determine high amplitude distortion and low aliasing effects. Instead, high values of c determine low amplitude distortion and high aliasing effects.

As to the relationship between the weighting coefficients $w(i,j;\varepsilon,\eta,c)$ and the relative squared error $E(i,j;\varepsilon,\eta,c)$, two different strategies are proposed, called *hard* and *soft,* respectively. The *hard* strategy represents the natural choice in absence of noise because it assigns to $\boldsymbol{Z}(i,j)$ the value of $\hat{\boldsymbol{Z}}(i,j;\varepsilon,\eta,c)$ that corresponds to the minimum of the relative squared error. In this case, the weighting coefficients are binary variables defined as

$$w(i,j;\varepsilon,\eta,c) = \begin{cases} 1 & E(i,j;\varepsilon,\eta,c) = \min\limits_{\bar{\varepsilon},\bar{\eta},\bar{c}} E(i,j;\bar{\varepsilon},\bar{\eta},\bar{c}) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where, according to (8), $|\bar{\varepsilon}| \le \varepsilon_{\mathrm{UB}}$, $|\bar{\eta}| \le \eta_{\mathrm{UB}}$, and $c \in [c_{\mathrm{LB}}, c_{\mathrm{UB}}]$. The resulting algorithm is referred to as RSFIM-HP-*hard* and is expected to effectively reproduce both the edges and the homogeneous regions in the super-resolved image $\boldsymbol{Z}(i,j)$.

In the presence of noise, which in each pixel can be realistically modeled as a spectrally uncorrelated zero mean random vector [50], [51], the *hard* strategy tends to bring the noise components of the HS image in the super-resolved image. To mitigate noise effects, the *soft* strategy is proposed where the weighting coefficients $w(i,j;\varepsilon,\eta,c)$ are defined so as to assign greater weights to solutions more consistent with the higher spatial resolution image, i.e., corresponding to low values of $E(i,j;\varepsilon,\eta,c)$:

$$w(i,j;\varepsilon,\eta,c) = \alpha(i,j)\, e^{-0.5\cdot E(i,j;\varepsilon,\eta,c)}. \tag{10}$$

In (10), $\alpha(i,j)$ is the normalization factor that makes the sum of the weights in each pixel equal to one. The resulting algorithm is referred to as RSFIM-HP-*soft*.

RSFIM-HP-*soft* is expected to be more robust to noise thanks to the filtering effect introduced by the weighted sum. This is particularly true in the homogeneous regions of the scene where the weights are very similar. Conversely, it is expected to introduce a smoothing effect on sharp variations thus performing worse than RSIM-HP-*hard* in preserving edges. In the experimental part of the article, we will give evidence of the previous comments.

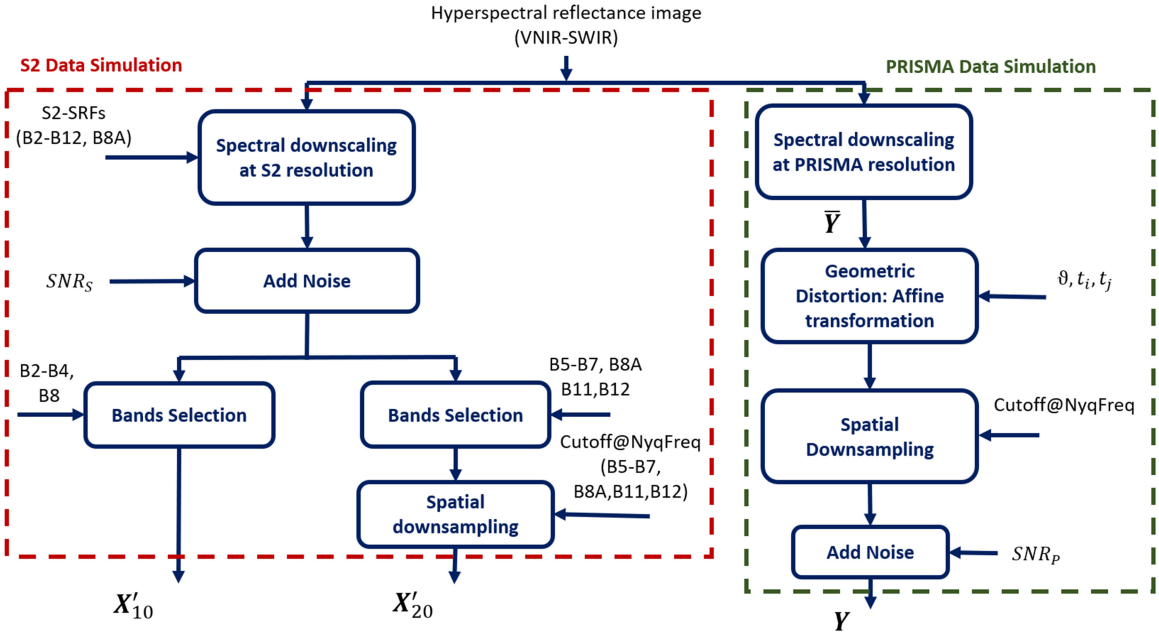


Fig. 4. Bock diagram of the adopted simulation strategy.

Fig. 5. RGB of the images adopted in our simulations: (a) AVIRIS-NG Grosseto and (b) RIT-Avon.

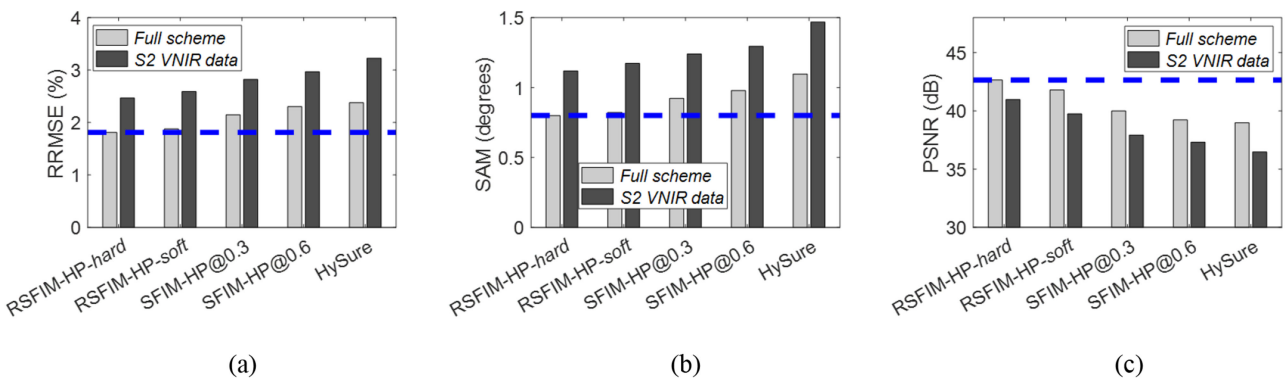


Fig. 6. AVIRIS-NG Grosseto, performance comparison between the full scheme and the S2 VNIR data experiments: (a) average RRMSE values, (b) average SAM values, and (c) average PSNR values. In all the figures, the dashed lines indicate the best performance.

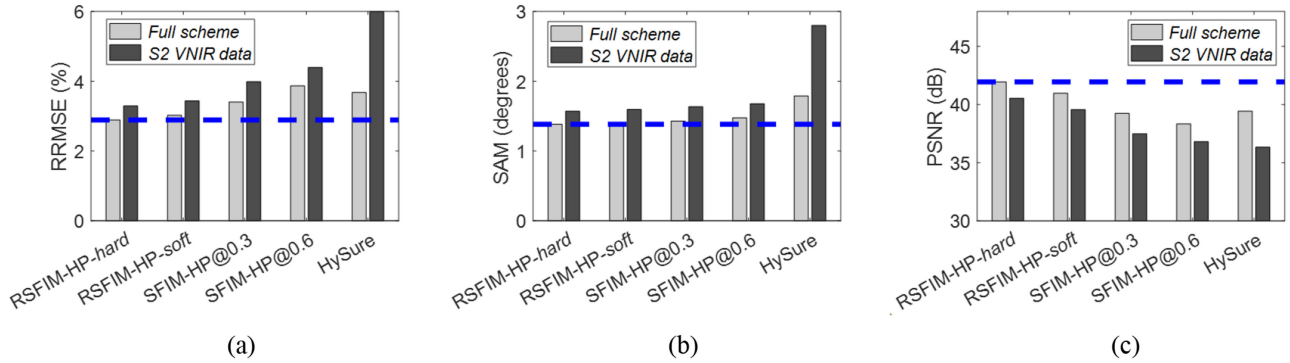


Fig. 7. RIT-Avon, performance comparison between the full scheme and the S2 VNIR data experiments: (a) average RRMSE values, (b) average SAM values, and (c) average PSNR values. In all the figures, the dashed lines indicate the best performance.

The two RSFIM-HP algorithms are implemented adopting a greed strategy that computes the quantities in (6)–(10) for discrete values of the parameters ε, η, and c within the assigned ranges.

The entire PRISMA-SR scheme with the algorithms proposed for each step and the symbols adopted in this section is shown in Fig. 3.

IV. Analysis on Simulated Data

In this section, we test the proposed end-to-end PRISMA-SR procedure on simulated data. We start by describing the simulation strategy adopted. Then, we present the datasets for simulation and the indices for performance evaluation. Finally, we present and discuss the experimental results.

Throughout the article, we consider reflectance data, i.e., the L2 product for both PRISMA and S2 images.

A. Simulation Strategy

The strategy adopted to simulate S2 and PRISMA data is depicted in Fig. 4. It starts from a given HS image acquired by a sensor that covers the VNIR–SWIR spectral range corresponding to the PRISMA data. The scheme in Fig. 4 branches into two parts. The left one refers to the simulation of the S2 data, while the right branch concerns the simulation of the PRISMA data.

S2 simulation branch first spectrally downsamples input data so as to simulate the spectral data in 10 S2 bands: B2–B12 and B8A. For this purpose, the SRFs provided by ESA [36] are adopted. Then, Gaussian random noise is added to each band. Noise power is set to obtain a predefined signal-to-noise ratio (SNR_S). Finally, the bands corresponding to B2–B4 and B8 are selected to simulate the full spatial resolution S2 image $\boldsymbol{X}'_{10}$. The remaining bands are downscaled of a factor of 2 in both the spatial directions to obtain the lower resolution S2 image $\boldsymbol{X}'_{20}$.

The right branch first derives the full spatial resolution PRISMA image $\bar{\boldsymbol{Y}}$ adopted as reference for performance assessment of the data fusion scheme. $\bar{\boldsymbol{Y}}$ is obtained by spectrally downsampling the input data to the spectral resolution of the PRISMA sensor. For this purpose, a Gaussian shaped SRF is considered for each band, with the central wavelength and the full width half maximum set according to the corresponding data provided by ASI.

$\bar{\boldsymbol{Y}}$ is spatially distorted by applying an affine transformation defined by the rotation angle ϑ and the shifts t_i and t_j (pixel units) in the two spatial directions.

The obtained image is spatially downsampled with a factor of 3 in both the directions to simulate the low spatial resolution PRISMA image. In this phase, a Gaussian low-pass filter is used to simulate the sensor MTF. The value of amplitude response of the filter at the Nyquist frequency for all the PRISMA sensor channels is set to 0.3 according to the experimental results in [52]. Finally, Gaussian noise with a given SNR (SNR_P) is added to the resulting image to obtain the simulated PRISMA image $\boldsymbol{Y}$.

B. Datasets

In our simulations, we use the two reflectance images described here below in more details. The red–green–blue (RGB) representation of the two images are shown in Fig. 5.

1) AVIRIS-NG Grosseto: this image is part of the publicly available data[2] acquired by the AVIRIS—Next Generation (NG) sensor on 7 April, 2018, nearby the city of Grosseto, Italy. AVIRIS-NG is an airborne sensor collecting data in 480 bands in the VNIR–SWIR range and with spectral sampling of 5 nm. The original data have a GSD of about 4 m. To get an image with GSD as close as possible to that of the higher resolution S2 bands, they have been downscaled by a factor of two in both the spatial directions. The resulting image has 264 $\times$ 600 pixels and a spatial resolution of about 8 m.
2) RIT-Avon: this HS image is part of the publicly available SHARE 2012 dataset[3] from the Digital Imaging and Remote sensing Laboratory of the Rochester Institute of Technology [53]. Specifically, we refer to the image acquired on 20 September, 2012 at 15:46 (GMT), in a rural scene located in AVON, New York, just south of Rochester. The image was collected by ProSpecTIR VS system, an HS airborne pushbroom spectrometer operating in the VNIR and SWIR spectral ranges and having 360 spectral bands and a spatial resolution lower than 1 m. In this

[2][Online]. Available: https://avirisng.jpl.nasa.gov/dataportal/
[3][Online]. Available: https://www.rit.edu/cos/share2012/

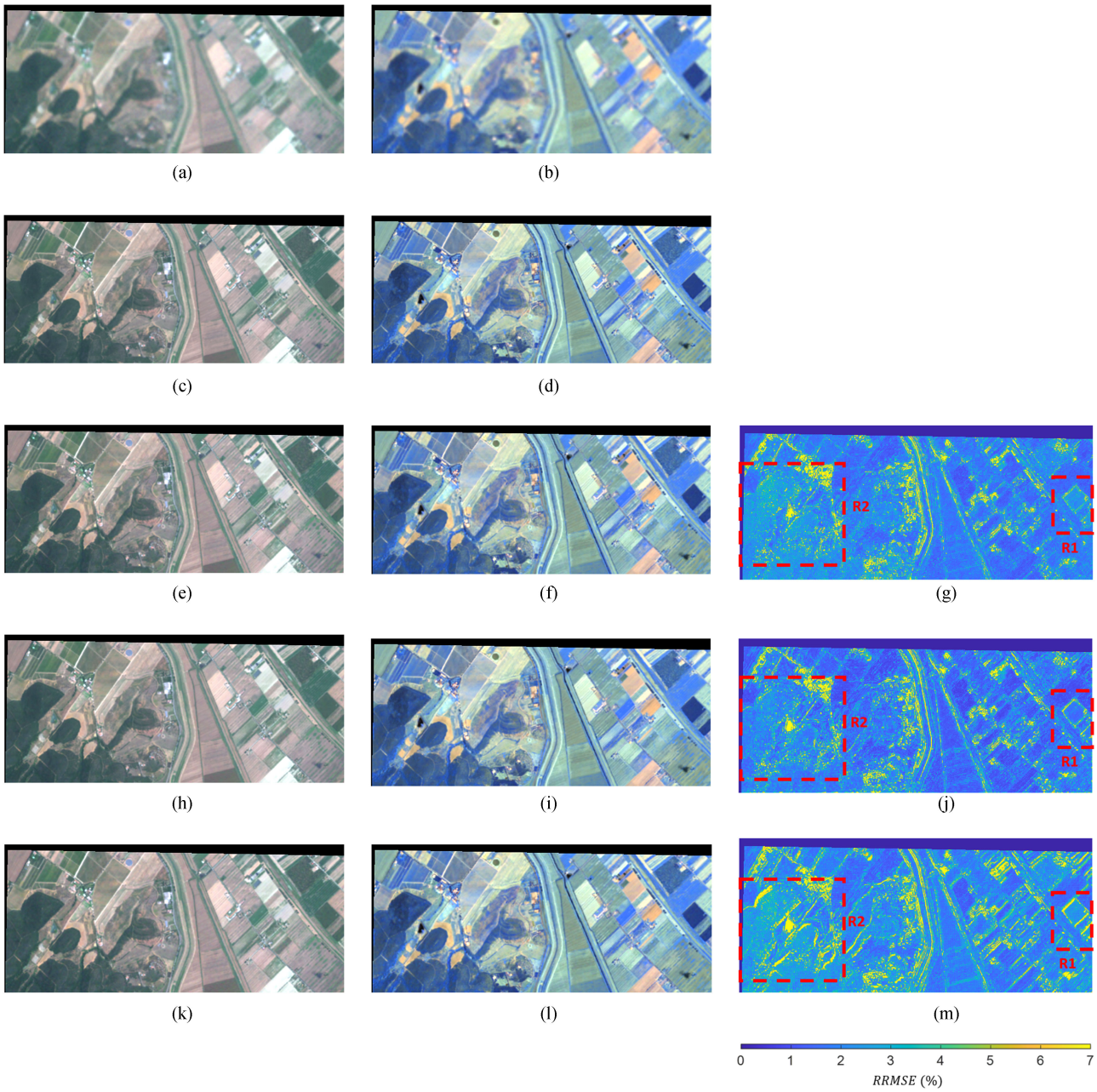


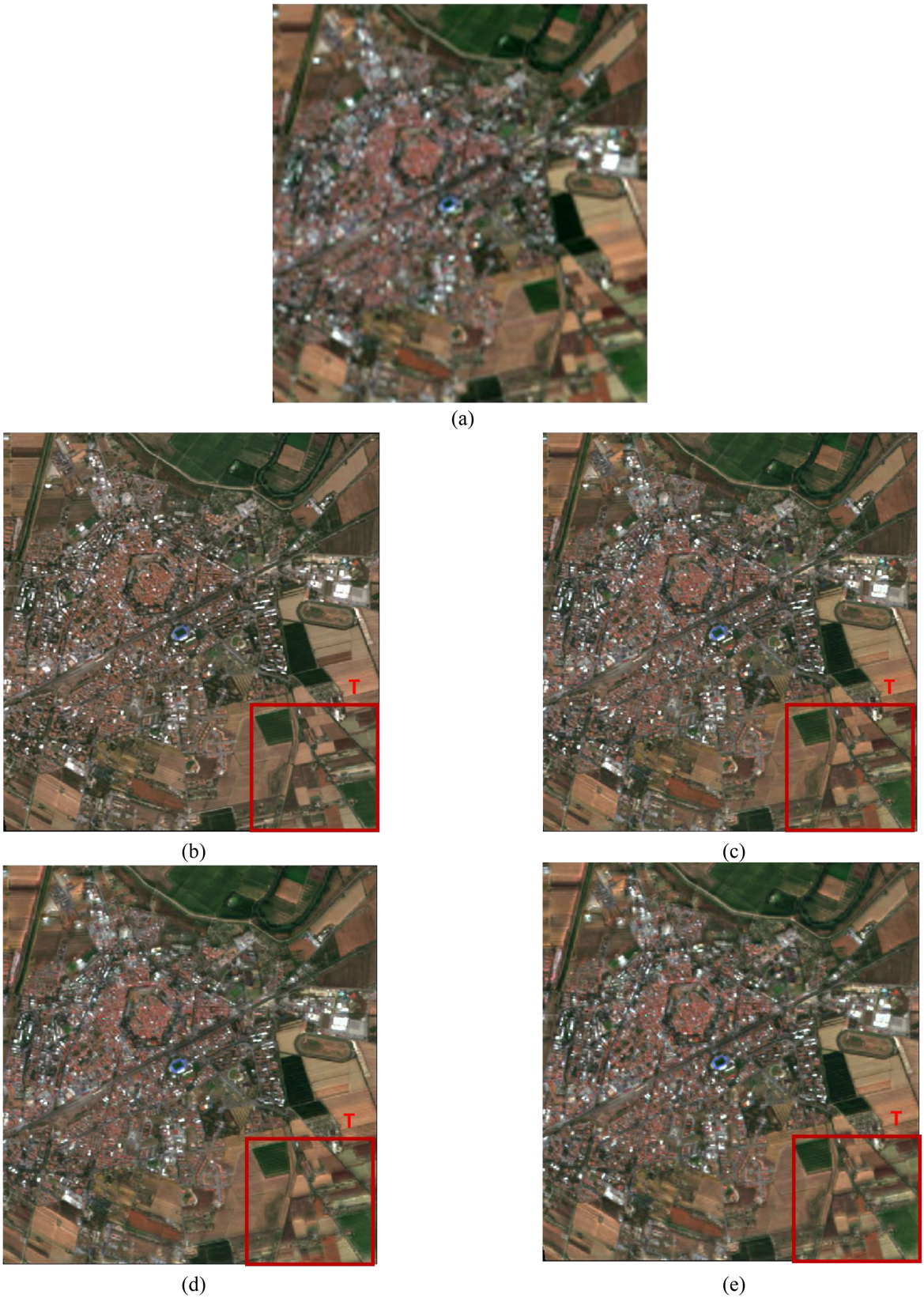
Fig. 8. Example of the results of the PRISMA-SR scheme on the AVIRIS-NG Grosseto dataset obtained with $\vartheta = 1$, $t_i = t_j = 1$, and $\mathrm{SNR}_P = 35$ dB: (a) RGB of the PRISMA upsampled image; (b) false color representation of the PRISMA upsampled image; (c) RGB of the PRISMA reference image; (d) false color representation of the PRISMA reference image; (e) RGB of the RSFIM-HP-hard output; (f) false color representation of the RSFIM-HP-hard output; (g) RRMSE (%) for the RSFIM-HP-hard algorithm; (h) RGB of the RSFIM-HP-soft output; (i) false color representation of the RSFIM-HP-soft output; (j) RRMSE (%) for the RSFIM-HP-soft algorithm; (k) RGB of the SFIM-HP output; (l) false color representation of the SFIM-HP output; and (m) RRMSE (%) for the SFIM-HP algorithm.

case, considering that the swath width is about 320 m, the full-resolution data have been used to have a sufficient number of pixels for testing the PRISMA-SR scheme. The total number of pixels in the adopted image is 264×600.

The high spatial/spectral resolution images used in our simulations have been acquired by airborne sensors. The high spatial resolution and the limited swath prevented us from simulating pixels at the true PRISMA spatial resolution. The resulting number of simulated pixels would have been too low for any subsequent analysis. For this reason, we have simulated data by maintaining the relative GSD ratios for the $\boldsymbol{X}'_{10}$, $\boldsymbol{X}'_{20}$, and $\boldsymbol{Y}$ images rather than imposing the real value of the GSD.

C. Performance Indexes

To assess the performance of the PRISMA-SR fusion scheme, we adopted three widely used quality measures: spectral angle mapper (SAM), peak signal-to-noise ratio (PSNR), and relative root mean squared error (RRMSE). Note that the PRISMA-SR fusion output $\boldsymbol{Z}$ is obtained in the coordinate reference system of



(a) (b) (c) (d) (e)

Fig. 9. Result on real data. RGB representations of (a) PRISMA upsampled; (b) Sentinel 2; (c) RSFIM-HP-hard; (d) RSFIM-HP-soft; and (e) SFIM-HP.

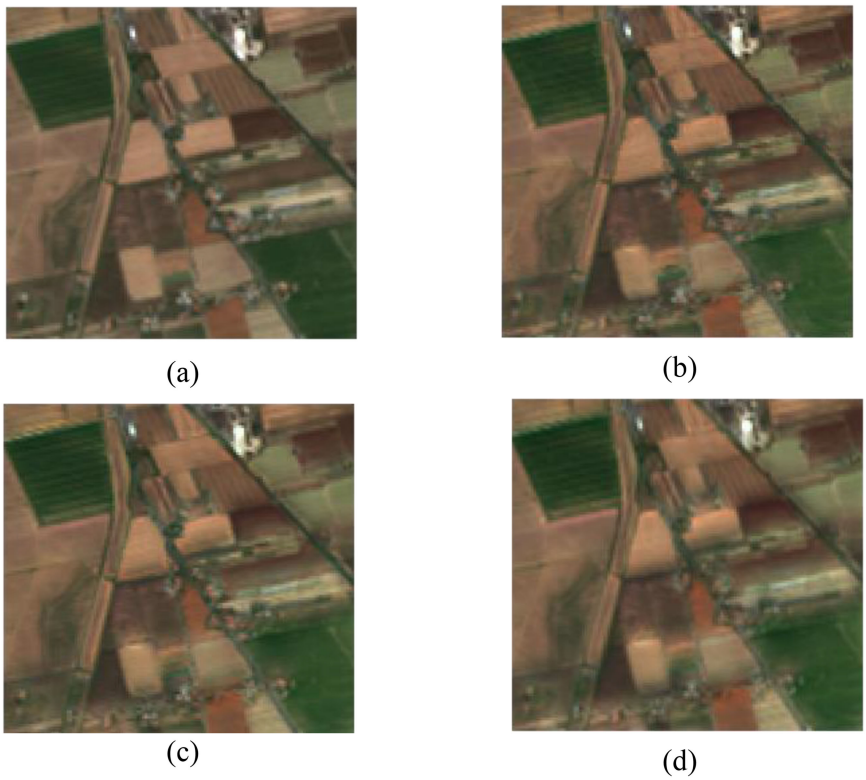


Fig. 10. RGB representations of the region of interest denoted as *T*: (a) Sentinel 2; (b) RSFIM-HP-hard; (c) RSFIM-HP-soft; and (d) SFIM-HP.

the S2 sharpened image after coregistration ($\boldsymbol{X}$). Thus, all the performance indexes are computed by assuming as reference the high spatial resolution PRISMA image $\boldsymbol{Y}_{\mathrm{REF}}$ obtained by applying to $\bar{\boldsymbol{Y}}$ (Fig. 4) the geometrical transformation that returns $\boldsymbol{X}$.

SAM measures the spectral similarity between each pixel of the estimated super-resolved image and the reference one:

$$\mathrm{SAM}\,(i,j) = \mathrm{acos}\left(\frac{\boldsymbol{Z}^{\boldsymbol{T}}\,(i,j)\cdot\boldsymbol{Y}_{\mathrm{REF}}\,(i,j)}{\|\boldsymbol{Z}\,(i,j)\,\|\cdot\|\boldsymbol{Y}_{\mathrm{REF}}\,(i,j)\,\|}\right). \quad (11)$$

The average value of SAM with respect to all the pixels is used as a performance index for the entire image. Values of SAM close to zero indicate high spectral quality.

RRMSE measures the relative error between each pixel of $\boldsymbol{Z}$ and $\boldsymbol{Y}_{\mathrm{REF}}$ as

$$\mathrm{RRMSE}\,(i,j) = \frac{\|\boldsymbol{Z}\,(i,j) - \boldsymbol{Y}_{\mathrm{REF}}\,(i,j)\,\|}{\|\boldsymbol{Y}_{\mathrm{REF}}\,(i,j)\,\|}. \quad (12)$$

Also, in this case, the average value with respect to all the pixels is used to evaluate the performance on the entire image.

PSNR is a measure of the spatial reconstruction quality, it is computed in the generic band b as

$$\mathrm{PSNR}\,(b) = 20\log_{10}\frac{\max\left(Y_{\mathrm{REF},b}\right)}{\|Z_b - Y_{\mathrm{REF},b}\|/\sqrt{N}} \quad (13)$$

where the maximum is taken on all the pixels of the image. The value of PSNR averaged over all the spectral bands is used as a synthetic index. High PSNR values denote good reconstruction quality.

D. Experimental Setup

In the experiments presented in this section, the three SFIM versions described in Section III-C are tested for the HS–MS fusion step. As to the original version of SFIM-HP, c is the only parameter to be set. Two values are considered $c = 0.3$ and $c = 0.6$. The first choice corresponds to the value adopted to simulate the MTF of the PRISMA HS sensor. The second one is introduced to show the sensitivity of the algorithm to the incorrect setting of the sensor MTF. The two robust versions of SFIM have three free parameters that define the range of variation of the local residual registration errors ($\varepsilon_{\mathrm{UB}}$ and η_{UB}) and the uncertainty about the value of the MTF at the Nyquist frequency (c_{LB} and c_{UB}). In all the experiments discussed below, we set $\varepsilon_{\mathrm{UB}} = \eta_{\mathrm{UB}} = 2$, $c_{\mathrm{LB}} = 0.2$, and $c_{\mathrm{UB}} = 0.7$.

As a further element of comparison for the HS–MS fusion step, we consider the state-of-the-art algorithm named HySure [24]. Specifically, we adopt the Matlab implementation of the algorithm[4] and the parameters setting presented in [14], on the basis of the stable and competitive performance obtained over different datasets.

As concerns the coregistration block of the PRISMA-SR scheme, we adopt the SURF and the image interpolation functions included in the Matlab Image Processing Toolbox.

[4]The source code of HySure is included in the HSMSFusionToolbox available at https://naotoyokoya.com/Publications.html.

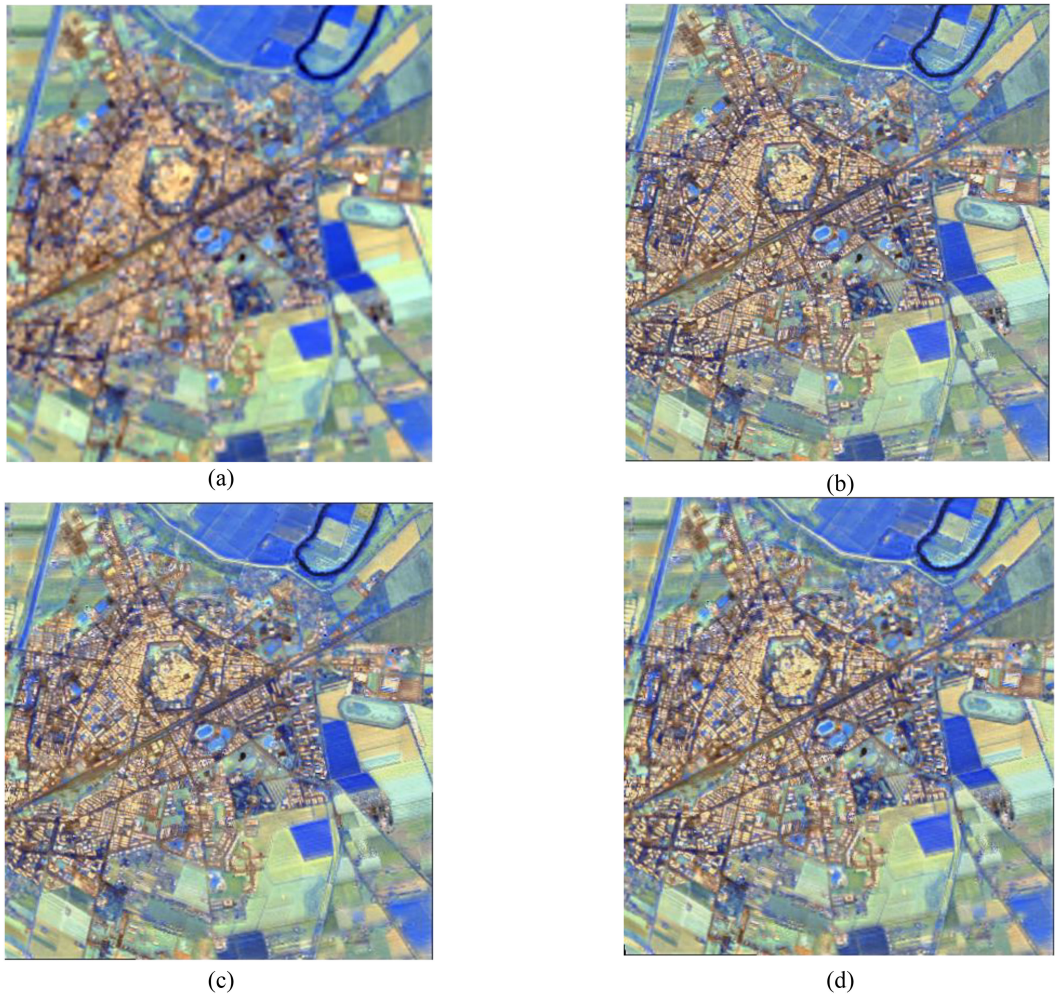


Fig. 11. Result on real data. False color composite images obtained by considering the PRISMA channels closest to 1000, 1650, and 2140 nm. (a) PRISMA upsampled. (b) Sentinel 2. (c) RSFIM-HP-hard. (c) RSFIM-HP-soft. (d) SFIM-HP.

E. Results

For each dataset described in Section IV-B and according to the simulation strategy in Section IV-A, $N = 50$ triplets of images $\boldsymbol{X}'_{10}$, $\boldsymbol{X}'_{20}$, and $\boldsymbol{Y}$ are generated by varying the noise realizations (for given values of SNR_S and SNR_P) and the geometric distortion parameters ϑ , t_i, and t_j. Specifically, values of ϑ in $[-2, 2]$ (°), t_i and t_j in $[-5, 5]$ (pixel units) are considered. The algorithms are applied to all the generated data and the values of the performance indexes averaged over all the realizations are adopted for the discussion. In all the experiments, we set $\mathrm{SNR}_S = 45$ dB and $\mathrm{SNR}_P \leq \mathrm{SNR}_S$.

Four types of experiments are performed to test the impact of the various processing steps of the proposed PRISMA-SR scheme.

The first one, called *Full scheme* hereinafter, consists in applying the entire PRISMA-SR scheme in Fig. 3.

The second experiment evaluates the impact of the SURF based coregistration block. To this end, the coregistration block is bypassed, and $\boldsymbol{X}$ is obtained from $\boldsymbol{X'}$ using the same geometrical transformation adopted to generate the PRISMA simulated image $\boldsymbol{Y}$. This experiment will be referred to as *Perfect coregistration*, in the rest of the article.

The third experiment investigates the impact of the S2 data fusion block on the performance of the complete PRISMA-SR scheme. For this purpose, the S2 data fusion block in Fig. 3 is bypassed, and the simulated full spatial resolution S2 image in the 10 bands B2–B8, B8A, B11, and B12 is provided as input to the SURF based coregistration block in place of $\boldsymbol{X'}$. It will be referred to as *Ideal S2 Sharpening*, hereinafter.

The fourth experiment, named *S2 VNIR data*, aims at investigating the benefits of the fusion of the S2 data at different spatial resolution compared to the case of adopting only the four S2 bands (B2–B4 and B8) in the VNIR spectral range with 10 m spatial resolution. This is carried out bypassing the S2 data fusion block and providing $\boldsymbol{X}'_{10}$ as input to the SURF-based coregistration block in place of $\boldsymbol{X'}$.

We start by discussing the results obtained in the case of high SNR for the PRISMA image. To this end, we set $\mathrm{SNR}_P = 45$ dB . Tables III and IV show the average values of the three performance indexes obtained in the *Full scheme*, the *Perfect coregistration*, and the *Ideal S2 Sharpening* experiments. Specifically, Tables III and IV refer to the AVIRIS-NG Grosseto dataset and the RIT-Avon dataset, respectively. In both the tables,

TABLE III
AVIRIS-NG GROSSETO: VALUES OF THE PERFORMANCE INDEXES OBTAINED IN THE CASE OF $\mathrm{SNR}_P = 45$ dB

	Full scheme			Perfect co-registration			Ideal S2 sharpening		
	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)
RSFIM-HP-*hard*	**1.81**	**0.79**	**42.64**	1.76	0.78	43.13	**0.87**	**0.44**	**49.20**
RSFIM-HP-*soft*	1.87	0.82	41.78	1.81	0.79	42.18	1.12	0.54	45.80
SFIM-HP $c = 0.3$	2.14	0.92	39.99	**1.70**	**0.76**	**43.77**	1.50	0.70	44.01
SFIM-HP $c = 0.6$	2.30	0.97	39.22	1.87	0.81	41.80	1.73	0.79	42
HySure	2.37	1.09	38.89	2.35	1.01	39.37	1.64	0.78	41.22

TABLE IV
RIT-AVON: VALUES OF THE PERFORMANCE INDEXES OBTAINED IN THE CASE OF $\mathrm{SNR}_P = 45$ dB

	Full scheme			Perfect co-registration			Ideal S2 sharpening		
	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)
RSFIM-HP-*hard*	**2.89**	**1.38**	**41.94**	2.84	1.38	42.14	**1.86**	**0.97**	**46.75**
RSFIM-HP-*soft*	3.02	**1.38**	40.97	2.95	1.38	41.20	2.1	1.01	44.61
SFIM-HP $c = 0.3$	3.41	1.42	39.24	**2.80**	**1.35**	**42.32**	2.88	1.08	41.85
SFIM-HP $c = 0.6$	3.86	1.47	38.33	3.31	1.40	40.21	3.39	1.15	40.42
HySure	3.67	1.78	39.43	3.64	1.77	39.83	3.64	1.48	41.94

we indicate in bold the best values of the performance indexes and underline the worst values. Several conclusions can be drawn from the observation of the numerical values:

1) Focusing on the *Full scheme* experiment, we can see that, for both the datasets, the procedures that adopt the two proposed versions of the RSFIM-HP algorithm provide the best performance. This is a first evidence of the effectiveness of the automatic compensation for both the residual coregistration errors and the uncertainty on the c value. In this case, comparing the two RSFIM-HP, we note that RSFIM-HP-*hard* performs the best. Furthermore, the original SFIM-HP achieves the best performance when the correct c value ($c = 0.3$) is adopted, thus, highlighting the sensitivity of the method to the c parameter. It is worth noting that all the tested methods provided pretty good performance. For example, we obtained RRMSE values lower than 2.4% in the *AVIRIS-NG Grosseto* dataset and lower than 3.9% in the *RIT-Avon* dataset.
2) Analyzing the performance in the case of *Perfect coregistration*, we can see that in both the datasets the best performance is achieved by the SFIM-HP fusion scheme with the correct value of c. Comparison of the performance indexes for SFIM-HP ($c = 0.3$) obtained in this case and in the previous discussed case, highlights the sensitivity of the algorithm to residual coregistration errors. For a numerical example, the RRMSE grows from 1.70% to 2.14% for the AVIRIS-NG Grosseto and from 2.8% to 3.41% for the RIT-Avon, in the presence of residual coregistration errors. The two RSFIM-HP algorithms provide values of the indexes very close to the best ones and, above all, very close to those obtained in the presence of residual coregistration errors. This is an experimental evidence of the robustness of the two algorithms to that type of errors. Here again, we can note that both the RSFIM-HP algorithms perform better than SFIM-HP with the wrong value of c, thus highlighting the effectiveness of the compensation mechanism for the uncertainty of such parameter.
3) As to the *Ideal S2 sharpening* experiment, it can be seen that RSFIM-HP-*hard* is the best performing algorithm. Comparing the performance of each algorithm with that obtained in the *Full scheme* case, we conclude that the impact of the S2 data fusion algorithm is not negligible. This encourages further studies and analysis aimed at investigating the possibility of improving the performance of the PRISMA-SR procedure by employing more

TABLE V
AVIRIS-NG GROSSETO: PERFORMANCE OF THE PRISMA-SR SCHEME FOR THREE SNR_p VALUES

	$SNR_p = 25\ dB$			$SNR_p = 35\ dB$			$SNR_p = 45\ dB$		
	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)
RSFIM-HP-*hard*	4.79	2.7	34.87	2.41	1.21	40.09	**1.81**	**0.79**	**42.64**
RSFIM-HP-*soft*	3.72	2.02	36.69	**2.21**	**1.06**	**40.41**	1.87	0.82	41.78
SFIM-HP $c = 0.3$	5.22	2.91	34.06	2.70	1.35	38.31	2.14	0.92	39.99
HySure	**3.22**	**1.66**	**37.04**	2.40	1.13	38.66	2.37	1.09	38.89

sophisticated and effective algorithms in the S2 data fusion step. However, the idea proposed in this article of using S2 sharpened data in the VNIR–SWIR spectral range, is still a good solution to improve the spatial resolution of PRISMA data. To give an evidence of such an improvement, in Figs. 6 and 7, we compare the performance obtained in the *Full scheme* experiment with that obtained when only the four pure full resolution S2 bands are adopted (*S2 VNIR data*). Specifically, Figs. 6 and 7 show the values of the performance indexes obtained on the AVIRIS-NG Grosseto dataset and the RIT-Avon dataset, respectively. For completeness, for each index, we have pointed out the best performance (dashed line). Figs. 6(a)–(b) and 7(a)–(b) show that, regardless of the HS–MS fusion algorithm adopted, both RRMSE and SAM values increase when the pure full resolution S2 bands (B2–B4 and B8) are used in place of the 10 bands S2 data fusion product. Similarly, Figs. 6(c) and 7(c) show that PSNR decreases when only the pure S2 bands are considered.

To analyze the impact of noise in the PRISMA data on the performance of the proposed PRISMA-SR scheme (*Full scheme*), experiments by varying SNR_P have been performed. Specifically, in addition to the high SNR case already discussed, we have considered the case of moderate noise ($\mathrm{SNR}_P = 35\,\mathrm{dB}$) and that of low SNR ($\mathrm{SNR}_P = 25\ \mathrm{dB}$).

Table V, with reference to the AVIRIS-NG Grosseto dataset and for the three considered values of SNR_P, shows the values of the performance indexes for each HS–MS fusion algorithm tested for the PRISMA-SR scheme. As to the SFIM-HP algorithm, only the case of correct value of the c parameter is considered. For completeness, Table V also shows the results obtained with $\mathrm{SNR}_P = 45\,\mathrm{dB}$ already discussed in the previous part of this Section. Notice that, in Table V (and subsequently in Table VI) we indicate in bold the best values of the performance indexes and underline the worst values. Table V shows that, with $\mathrm{SNR}_P = 35\ \mathrm{dB}$, RSFIM-HP-*soft* is the best performing algorithm. This result, compared with the case of high SNR, gives an experimental evidence of the better robustness to noise of the *soft* version of RSFIM-HP with respect to the *hard* one. As stated in Section III-C, the weighting strategy adopted in the *soft* version reduces the variance of spatially uncorrelated noise and the sensitivity of the output to the noise effect. Such a property is confirmed by the comparison of the results obtained by applying the two versions of RSFIM-HP in the low SNR case. SFIM-HP shows a greater sensitivity to noise.

As to HySure, results in Table V, show that it is the best performing HS–MS fusion algorithm in the case of low SNR.

The different in noise sensitivity between HySure and the SFIM-based algorithms (RSFIM-HP in both the proposed versions and SFIM-HP) was not unexpected. In fact, according to (4) and (6), in SFIM-based algorithms the super-resolved image in each band is directly proportional to the upsampled version of the low-resolution HS image $\boldsymbol{Y}$. So the noise contribution in $\boldsymbol{Y}$ is transferred to the $\boldsymbol{Z}$ output. As remarked in Section III-C, a form of noise mitigation is obtained only in the RSFIM-HP soft as a consequence of the averaging operation introduced by the weighting sum of the various contributions.

On the contrary, HySure benefits from the noise mitigation introduced by both the subspace projection related to the adopted linear mixture model and the total variation regularization used to address the ill-posedness of the approached optimization problem [14].

It is worth remarking that in the cases of high and moderate SNR, the detrimental effect of the residual coregistration error exceeds that of noise, and the proposed RSFIM-HP algorithms perform better than HySure.

Very similar conclusions can be drawn from the results obtained for the RIT-Avon dataset, reported in Table VI.

As an example, in Fig. 8(a)–(m), we show the results obtained by applying the PRISMA-SR scheme to one of the triplets of images $\boldsymbol{X}'_{10}$, $\boldsymbol{X}'_{20}$, and $\boldsymbol{Y}$ simulated from the AVIRIS-NG Grosseto data. Only the results obtained by the two proposed RSFIM-HP algorithms and the SFIM-HP ($c = 0.3$) algorithm are shown in Fig. 8. Specifically, they refer to the case of $\vartheta = 1$, $t_i = t_j = 1$, and $\mathrm{SNR}_P = 35\ \mathrm{dB}$.

Fig. 8(a) and (b) shows the RGB representation and the false color (FC) representation of the simulated PRISMA image upsampled by a factor of 3, respectively. The FC representation is obtained by considering the three PRISMA channels with the central wavelengths closest to 1000, 1650, and 2140

TABLE VI
RIT-AVON: PERFORMANCE OF THE PRISMA-SR SCHEME FOR THREE SNR_p VALUES

	$SNR_p = 25\ dB$			$SNR_p = 35\ dB$			$SNR_p = 45\ dB$		
	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)	RRMSE (%)	SAM (degrees)	PSNR (dB)
RSFIM-HP-*hard*	6.42	3.60	37.82	3.56	1.81	**40.92**	**2.89**	**1.38**	**41.94**
RSFIM-HP-*soft*	5.21	2.75	**38.64**	**3.42**	**1.62**	40.66	3.02	**1.38**	40.97
SFIM-HP $c = 0.3$	6.91	3.77	36.73	3.92	1.85	39.59	3.41	1.42	39.24
HySure	**4.86**	**2.52**	37.73	3.71	1.71	39.17	3.67	1.78	39.43

TABLE VII
COMPUTATION TIME FOR EACH STEP OF THE PRISMA-SR SCHEME

Processing step		CPU time (sec)
S2 data fusion: MTF-GLP-CBD		0.21
SURF-based co-registration		2.66
HS-MS Fusion		
	RSFIM-HP-*hard*	34.70
	RSFIM-HP-*soft*	34.70
	SFIM-HP	2.63
	HySure	241.45

nm. They have been chosen by considering the portion of the spectrum not covered by the S2 data and outside the strongest atmospheric absorption windows. Such an FC representation gives an example of the SR results in the spectral range not covered by the S2 sensor. Fig. 8(c) and (d) shows the RGB and the FC representations of the full-resolution PRISMA image (the reference). The RGB and the FC images are also shown for the outputs of all the considered algorithms together with RRMSE per pixel (in percent). The visual comparison among the RGB and the FC images provided by each algorithm and that of the reference image shows that each of them provides good SR results. In general, the two versions of the RSFIM-HP algorithm provide RRMSE values lower than those obtained by the SFIM-HP algorithm. This is particularly evident in the two regions marked as R1 and R2 in Fig. 8(g), (j), and (m). Comparing the RRMSE of the two RSFIM-HP algorithms in the R1 region, it is quite evident that the *hard* version has a better ability of preserving the edges in the scene. Conversely, the analysis of the RRMSE in the R2 region, shows that the *soft* strategy outperforms the *hard* one in the homogeneous regions of the scene due to its better ability to filter out spatially uncorrelated noise.

To conclude this section, we consider the computational time required by the proposed PRISMA-SR scheme. For this purpose, with reference to the AVIRIS-NG datasets, in Table VII, we report the computational time required by each step of the proposed scheme. As to the HS–MS fusion, we compare the two versions of RSFIM-HP algorithm, the SFIM-HP algorithm, and HySure. Results, in Table VII, have been obtained by using the Matlab implementation of each step of the processing chain on a personal computer equipped with Intel Core i9-11900 at 2.5 GHz and a RAM of 32 Gb.

Table VII shows that HS–MS fusion is the most computationally expensive processing step. As expected, the greedy strategy adopted to account for local registration errors and the uncertainty on the c parameter, makes the RSFIM-HP algorithms computationally more expensive than the original SFIM-HP. The computational time for RSFIM-HP is one order of magnitude higher than that of SFIM-HP. However, the computational time of the two versions of RSFIM-HP algorithm can be significantly reduced by taking advantage of its inherent parallel architecture. HySure, due to the iterative processing required by the ADMM optimization procedure, is the most computationally expensive algorithm.

IV. EXAMPLE ON REAL DATA

To give an example on real data, we have considered the PRISMA image acquired on 8 July 2020 over the city of Grosseto, Italy, and the closest (in time) data acquired by S2A on the same scene. Specifically, S2A data were acquired on 7 July 2020. For both PRISMA and S2A, we have considered the L2 data product.

A subimage of 140×130 pixels has been extracted from the PRISMA image and, based on the georeferencing information, properly overlapped 10 and 20 m S2 data have been selected. The proposed end-to-end PRISMA-SR procedure has been applied to the abovementioned data and the output obtained by adopting the RSFIM-HP-*hard*, RSFIM-HP-*soft*, and SFIM-HP are presented. As to SFIM-HP, the c parameter has been set to 0.3. For the two versions of RSFIM-HP, we have adopted the same settings used for the simulated data, i.e., $\varepsilon_{\mathrm{UB}} = \eta_{\mathrm{UB}} = 2$, $c_{\mathrm{LB}} = 0.2$, and $c_{\mathrm{UB}} = 0.7$.

Of course, since the high-resolution HS image is not available, we cannot perform any quantitative evaluation of the results.

Fig. 9(a) and (b) shows the RGB representations of the PRISMA image upsampled via bicubic interpolation so as to have a spatial resolution of 10 m, and the S2 image at the finest resolution (10 m) after coregistration with the PRISMA image, respectively. Fig. 9(c)–(e) shows the RGB representations of

the results provided by the proposed procedure for the three considered HS–MS fusion algorithms.

In all the cases considered, the PRISMA-SR procedure provides better results than bicubic interpolation and very similar to the high resolution S2 image. Fig. 10(a)–(d) shows the RGB representation of the region of the scene denoted as T in Fig. 9(b)–(e). The figures show that the images obtained by RSFIM-HP-*soft* [Fig. 10(c)] and by SFIM-HP [Fig. 10(d)] are blurred if compared to the full resolution S2 image [Fig. 10(a)]. Such blurring, probably due to the residual coregistration error, is less evident in the image obtained by RSFIM-HP-*hard* [Fig. 10(b)].

For completeness, in Fig. 11(a)–(d), we show the FC composite images corresponding to the PRISMA upsampled image and to the results obtained by the three considered fusion algorithms. As in Section IV, FC representations are obtained by considering three channels of the PRISMA sensor falling in the region of the spectrum not covered by S2. One can clearly see that results provided by the PRISMA-SR procedure are much more sharpen than those obtained by bicubic interpolation. Furthermore, given the color similarity among each super-resolved image and the bicubic interpolated image, PRISMA-SR procedure does not seem to introduce significant spectral distortions in the considered bands.

V. Conclusion

In this article, a new end-to-end procedure called PRISMA-SR has been presented to increase the spatial resolution of PRISMA HS data. The main idea behind the proposed procedure is that of exploiting S2 data acquired with spatial resolution of 10 and 20 m to obtain a super-resolved PRISMA image with a pixel size of 10 m. For this purpose, S2 data at different spatial resolutions are preliminarily fused to obtain a synthetic MS image with 10 m spatial resolution and 10 spectral bands. This fused image, together with the low-resolution PRISMA image, is adopted by the HS–MS fusion scheme to obtain the final PRISMA super-resolved data. Before performing the fusion task, the two images are coregistered in order to compensate georeferencing errors.

In the article, we have proposed a solution for each step of the PRISMA-SR procedure. Particularly, for the HS–MS fusion step, two modified versions (RSFIM-HP-*hard* and RSFIM-HP-*soft*) of the well-known SFIM-HP algorithm have been proposed. Such versions are specifically designed to account for residual errors that inevitably remain after the application of the coregistration algorithm. Furthermore, unlike SFIM-HP, the two RSFIM-HP algorithms automatically compensate for the uncertainty on the c parameter of the Gaussian kernel adopted to approximate the sensor MTF. Both the RSFIM-HP algorithms have three free parameters of pretty intuitive interpretation. Particularly, two of them can be set according to the expected local co-registration residual errors.

Several experiments on simulated data have been discussed to test the PRISMA-SR procedure. Results have shown its effectiveness in realistic scenarios.

Results have also shown the robustness of the proposed RSIM-HP algorithms to residual coregistration error and their ability to compensate for uncertainty on sensor MTF.

One of the main conclusions drawn from the presented analysis is that, exploiting the S2 data fusion for PRISMA SR is indeed effective. In fact, in all the experiments, we found that performance obtained by S2 data fusion is better than that obtained by using the four full resolution (10 m) S2 bands. However, we have also shown that the impact of the S2 data fusion algorithm on the final product is not negligible. Performance obtained by assuming the availability of ten S2 bands at full resolution is better than that obtained by using the output of S2 data fusion. The previous remarks encourage future efforts in testing various S2 data fusion procedures to search for the best performing one in the PRISMA SR framework.

It is worth pointing out that, unlike most of the HS–MS fusion methods proposed in the literature, the two RSFIM-HP algorithms used in the proposed HS–MS fusion scheme take into account the residual coregistration error. In the experimental part of this article, RSFIM-HP algorithms are compared with the well-known HySure, which is regarded as a benchmark in most of the published articles for its state-of-the-art performance. Results have shown that, for high and moderate noise levels, in the presence of residual coregistration error the proposed solutions provide better fusion performance than HySure.

An important point that will be addressed in future work concerns the complementary properties of the two versions of the RSFIM-HP algorithm. The hard version has greater edge-preserving ability while the soft version is more robust to uncorrelated noise. We are studying a possible strategy to exploit the strengths of the two versions of the RSFIM-HP algorithm proposed in this article. The basic idea is similar to that adopted in bilateral filtering and consists in weighing the outputs of the two algorithms in each pixel position on the basis of the degree of homogeneity of the surrounding region.

Another important open issue is related to the assumption that data acquired by S2 and PRISMA do not include changes in the composition of the monitored scene. The hypothesis is reasonable for both natural and artificial surfaces if the two images are acquired very close in time. However, S2 and PRISMA images acquired with a very low time difference may not be available. Therefore, to increase the application domain of the proposed PRISMA-SR approach, it is necessary to define specific strategies capable of taking into account changes occurred in the monitored scene.

References

[1] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*. Hoboken, NJ, USA: Wiley, 2007.

[2] J. Transon, R. d'Andrimont, A. Maugnard, and P. Defourny, "Survey of hyperspectral earth observation applications from space in the Sentinel-2 context," *Remote Sens.*, vol. 10, no. 2, pp. 1–32, Jan. 2018.

[3] F. D. Van der Meer *et al.*, "Multi- and hyperspectral geologic remote sensing: A review," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 14, no. 1, pp. 112–128, 2012.

[4] P. S. Thenkabail, M. K. Gumma, P. Teluguntla, and I. A. Mohammed, "Hyperspectral remote sensing of vegetation and agricultural crops," *Photogramm. Eng. Remote Sens.*, vol. 80, no. 4, pp. 697–709, 2014.

[5] E. Adam, E. O. Mutanga, and D. Rugege, "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review," *Wetland Ecol. Manage.*, vol. 18, pp. 281–296, 2010.

[6] A. Dekker, V. Brando, J. Anstee, N. Pinnel, and A. Held, "Preliminary assessment of the performance of hyperion in coastal waters. Cal/Val activities in Moreton Bay, Queensland, Australia," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 9–13, 2001, pp. 2665–2667.

[7] S. Pignatti *et al.*, "Development of algorithms and products for supporting the Italian hyperspectral PRISMA mission: The SAP4PRISMA project," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 127–130.

[8] S. Pignatti *et al.*, "The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2013, pp. 4558–4661.

[9] R. Guarini *et al.*, "Prisma hyperspectral mission products," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 179–182.

[10] R. Loizzo *et al.*, "Prisma mission status and perspective," *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 4503–4506.

[11] Accessed: Mar. 31, 2021. [Online]. Available: http://prisma-i.it/index.php/en/

[12] G. Vivone *et al.*, "A new benchmark based on recent advances in multispectral pansharpening," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 1, pp. 53–81, Oct. 2020.

[13] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper sharpening: A first approach on SIM-GA data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3008–3024, Jun. 2015.

[14] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.

[15] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[16] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing dor hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[17] R. Kawakami, J. Wright, Y. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuci, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE CVPR*, 2011, pp. 2329–2336.

[18] E. Wicoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1409–1413.

[19] C. Lanaras, E. Baltsacias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3586–3594.

[20] M. A. Veganzones, M. Simões, G. Licciardi, N. Yokoya, J. M. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 274–288, Jan. 2016.

[21] Q. Wei, J. M. B. Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[22] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multiband images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

[23] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, Nov. 2015.

[24] M. Simoes, J. B. Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[25] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[26] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neutal network," *Remote Sens.*, vol. 8, no. 7, pp. 1–22, Jul. 2016.

[27] J. Ma, W. Yu, C. Chen, P. Liang, X. Guo, and J. Jiang, "Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion," *Inf. Fusion*, vol. 62, pp. 110–120, May 2020.

[28] H. Zhang and J. Ma, "GTP-Net: A residual learning network based on gradient transformation prior for pansharpening," *ISPRS J. Photogramm. Remote Sens.*, vol. 172, pp. 223–239, Jan. 2021.

[29] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5345–5355, Nov. 2018.

[30] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu, "Multispectral and hyperspectral image fusion by MS/HS fusion net," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 1585–1594.

[31] X. Han, Y. Zheng, and Y. Chen, "Multi-level and multi-scale spatial and spectral fusion CNN for hyperspectral image super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop*, 2019, pp. 4330–4339.

[32] W. Wei, J. Nie, Y. Li, L. Zhang, and Y. Zhang, "Deep recursive network for hyperspectral image super-resolution," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1233–1244, 2020.

[33] R. Dian, S. Li, B. Sun, and A. Guo, "Recent advances and new guidelines on hyperspectral and multispectral image fusion," *Inf. Fusion*, vol. 69, pp. 40–51, May 2021.

[34] M. Drusch *et al.*, "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, May 2012.

[35] J. Li and B. Chen, "Global revisit interval analysis of Landsat-8-9 and Sentinel-2a-2b data for terrestrial monitoring," *Sensors*, vol. 20, no. 22, pp. 1–15, Nov. 2020.

[36] Accessed Jun. 14, 2021. [Online]. Available: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument

[37] Q. Wang, W. Shi, Z. Li, and P. M. Atkinson, "Fusion of Sentinel-2 images," *Remote Sens. Environ.*, vol. 187, pp. 241–252, Dec. 2016.

[38] A. Vaiopoulos and K. Karantzalos, "Pansharpening on the narrow VNIR and SWIR spectral bands of Sentinel-2," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 41, pp. 723–730, Jun. 2016.

[39] H. Park, J. choi, N. Park, and S. Choi, "Sharpening the VNIR and SWIR bands of Sentinel-2A imagery through modified selected and synthesized band schemes," *Remote Sens.*, vol. 9, no. 10, pp. 1–20, Oct. 2017.

[40] C. Paris, J. Bioucas-Dias, and L. Bruzzone, "A novel sharpening approach for superresolving multiresolution optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1545–1560, Sep. 2018.

[41] M. Gargiulo, A. Mazza, R. Gaetano, G. Ruello, and G. Scarpa, "Fast super-resolution od 20 m Sentinel-2 bands using convolutional neural networks," *Remote Sens.*, vol. 11, no. 22, pp. 1–18, Nov. 2019.

[42] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Lecture Notes Comput. Sci.*, vol. 3951, pp. 404–417, 2006.

[43] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Jan. 2000.

[44] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[45] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF tailored multiscale fusion of high-resolution MS and pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[46] Accessed Jun. 11, 2021. [Online]. Available: https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_L1C_Data_Quality_Report.pdf/6ad66f15-48ca-4e65-b304-59ef00b7f0e0?t=1623227693274

[47] R. Bouchiha and K. Besbes, "Automatic remote-sensing images registration using SURF," *Int. J. Comput. Theory Eng.*, vol. 5, no. 1, pp. 88–92, 2013.

[48] M. Scaioni, L. Barazzetti, and M. Gianinetto, "Multi-image robust alignement of medium resolution satellite imagery," *Remote Sens.*, vol. 10, no. 1969, pp. 1–25, Dec. 2018.

[49] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[50] T. Skauli, "Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing," *Opt. Exp.*, vol. 19, no. 14, pp. 13031–13046, Jun. 2011.

[51] N. Acito, M. Diani, and G. Corsini, "Signal-dependent noise modeling and model parameter estimation in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2957–2971, Aug. 2011.

[52] P. Coppo *et al.*, "Leonardo spaceborne infrared payloads fo Earth observation: SLSTRs for Copernicus Sentinel 3 and PRISMA hyperspectral camera for PRISMA satellite," *Appl. Opt.*, vol. 59, no. 23, pp. 6888–6901, Aug. 2020.

[53] A. Giannandrea *et al.*, "The SHARE 2012 data campaign," in *Proc. Algorithms Technol. Multispectral, Hyperspectral, Ultraspectral Imagery*, 2013, pp. 87430F–1–15.

Nicola Acito (Member, IEEE) received the laurea degree (cum laude) in telecommunication engineering and the Ph.D. degree in methods and technologies for environmental monitoring from the University of Pisa, Pisa, Italy, in 2001 and 2005, respectively.

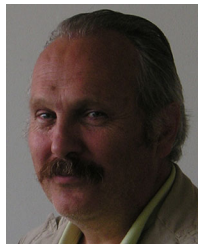
From November 2004 to October 2008, he was a Temporary Researcher with the Department of Information Engineering, University of Pisa. Form March 2009 to December 2020, he was a Researcher/Lecturer with the Italian Navy Academy. Since January 2021, he has been an Associate Professor with the University of Pisa. His research interests include multidimensional signal and image processing in remote sensing applications, deep learning, and machine learning techniques for remotely sensed data. His research interests also include atmospheric compensation and super-resolution of hyperspectral images.

Prof. Acito is an Associate Editor for the *Journal of Applied Remote Sensing* and has served as a Reviewer for many journals in the fields of remote sensing and image/signal processing, i.e., IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS.

Marco Diani (Member, IEEE) received the laurea degree (cum laude) in electronic engineering from University of Pisa, Pisa, Italy, in 1988.

He is currently a Full Professor with the Italian Naval Academy, Livorno, Italy, where he teaches "Signal Processing" and "Telecommunications Theory." He also teaches "Remote Sensing" with the University of Pisa. His research interests include image and signal processing with application to remote sensing. In this framework, his works have covered different topics such as data fusion, signal processing in imaging radars, image classification and segmentation, and object detection. His research interests also include development of new algorithms for object detection and tracking in infrared image sequences and on target detection and recognition in multi/hyperspectral images.

Prof. Diani is an Associate Editor for *the Journal of Applied Remote Sensing* and for the *Journal of Real Time Image Processing*. He has served as a Reviewer for many journals in the fields of remote sensing and image/signal processing such as IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and *Applied Optics*. He is member of SPIE.



Giovanni Corsini (Member, IEEE) received the laurea degree in electronic engineering from University of Pisa, Pisa, Italy, in 1979.

He is currently a Full Professor with the Department of Information Engineering, University of Pisa, where he teaches "Image Processing." His research interests include signal detection and processing with emphasis on image and multidimensional signal analysis in remote sensing applications. His research interests include object detection and parameter estimation from remotely sensed data with particular emphasis on hyperspectral and multispectral images.

Prof. Corsini has been a member of technical committees at international conferences. He is a Reviewer of the most important international journals in this field, i.e., IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, and *International Journal on Remote Sensing*.