



Article

Advantages of Nonlinear Intensity Components for Contrast-Based Multispectral Pansharpener

Alberto Arienzo ¹, Luciano Alparone ¹, Andrea Garzelli ^{2,*} and Simone Lolli ^{3,4}

¹ Department of Information Engineering, University of Florence, 50139 Florence, Italy; alberto.arienzo@unifi.it (A.A.); luciano.alparone@unifi.it (L.A.)

² Department of Information Engineering and Mathematics, University of Siena, 53100 Siena, Italy

³ CNR-IMAA, Consiglio Nazionale delle Ricerche, Contrada S. Loja snc, Tito Scalo, 85050 Potenza, Italy; simone.lolli@imaa.cnr.it or simone.lolli@upc.edu

⁴ CommSensLab, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

* Correspondence: andrea.garzelli@unisi.it

Abstract: In this study, we investigate whether a nonlinear intensity component can be beneficial for multispectral (MS) pansharpener based on component-substitution (CS). In classical CS methods, the intensity component is a linear combination of the spectral components and lies on a hyperplane in the vector space that contains the MS pixel values. Starting from the hyperspherical color space (HCS) fusion technique, we devise a novel method, in which the intensity component lies on a hyper-ellipsoidal surface instead of on a hyperspherical surface. The proposed method is insensitive to the format of the data, either floating-point spectral radiance values or fixed-point packed digital numbers (DNs), thanks to the use of a multivariate linear regression between the squares of the interpolated MS bands and the squared lowpass filtered Pan. The regression of squared MS, instead of the Euclidean radius used by HCS, makes the intensity component no longer lie on a hypersphere in the vector space of the MS samples, but on a hyperellipsoid. Furthermore, before the fusion is accomplished, the interpolated MS bands are corrected for atmospheric haze, in order to build a multiplicative injection model with approximately de-hazed components. Experiments on GeoEye-1 and WorldView-3 images show consistent advantages over the baseline HCS and a performance slightly superior to those of some of the most advanced methods.

Keywords: contrast-based image fusion; haze correction; multispectral pansharpener; multivariate regression; remote sensing



Citation: Arienzo, A.; Alparone, L.; Garzelli, A.; Lolli, S. Advantages of Nonlinear Intensity Components for Contrast-Based Multispectral Pansharpener. *Remote Sens.* **2022**, *14*, 3301.

<https://doi.org/10.3390/rs14143301>

Academic Editor: João Catalão Fernandes

Received: 1 June 2022

Accepted: 4 July 2022

Published: 8 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The availability of image data with spectral diversity (visible, near infrared, short wave infrared, thermal infrared, X- and C-band microwaves with related polarizations) and complementary spectral-spatial resolution, together with the peculiar characteristics of each image set, have fostered the development of fusion techniques specifically tailored to remotely sensed images of the Earth. Fusion aims at producing an extra value with respect to those separately available from the individual datasets. Though the results of fusion are more often analyzed by human experts to solve specific tasks (detection of landslides, flooded and burned areas, just to mention a few examples), partially supervised and also fully automated systems, most notably thematic classifiers, have started benefiting from fused images instead of separate datasets.

Extensive research on remote sensing image fusion for Earth observation has been carried out over the last decade, and a remarkable number of algorithms have been developed [1]. Image fusion techniques can be classified according to different criteria. One of the most common ways to differentiate fusion algorithms is based on sensor homogeneity. The term homogeneous image fusion refers to the case in which the images to be

merged are produced by sensors exploiting the same imaging mechanism. This category is also called *unimodal* image fusion. In remote sensing for Earth observation, the fusion of panchromatic and multispectral (MS) images, also known as pansharpening, is a typical example of homogeneous image fusion. The images subject to fusion are the outcome of measurements of the reflected solar radiation of the scene, even though they are referred to different wavelengths and are characterized by different information contents, also in terms of spatial resolution. On the other hand, the fusion of heterogeneous data, or *multimodal* image fusion, is referred to those cases in which the data to be merged come from sensors not sharing the same imaging mechanism.

An additional way to discriminate among fusion techniques is based on the content level subject to fusion, i.e., pixel level, feature level, and decision level [1]. Pixel level image fusion directly combines the pixels of the involved images in order to produce a new image, whereas feature level fusion aims to combine specific features or descriptors extracted from the images to be merged. The extraction of the features can be performed either simultaneously on all the images or separately on each image. As an example, for the fusion of optical and Synthetic Aperture Radar (SAR) images [2], a direct merge of the two datasets is not recommended, to prevent contamination of the fusion product with the low signal-to-noise ratio (SNR) of SAR data. In this case, features extracted from the SAR image, either [3] for texture and spatial heterogeneity or [4] for temporal coherence of the scene derived from geocoded multilooked products, can be transplanted into the optical image, thereby alleviating the stringent requirement of co-registration between the two datasets typical of pixel-level fusion. Decision level fusion is the combination of the classification results achieved either from each dataset separately or from multiple algorithms on the same dataset. In this case, the fusion output is a classification map [5].

Among pixel-based remote-sensing image-fusion techniques, panchromatic (Pan) sharpening, or pansharpening, of multispectral (MS) images is receiving ever increasing attention [1,6]. Pansharpening takes advantage of the complementary characteristics of the spatial and spectral resolutions of MS and Pan data, originated by physical constraints on the SNR of broad and narrow bands [7]. The goal is the synthesis of a unique product that exhibits as many spectral bands as the original MS image, each with same spatial resolution as the Pan image.

After the MS bands have been interpolated and co-registered to the Pan image [8], spatial details are extracted from Pan and added to the MS bands according to a predefined injection model. The detail extraction step may follow the *spectral* approach, originally known as component substitution (CS), or the *spatial* approach, which may rely on multiresolution analysis (MRA), either separable or not [9]. In the spectral approach, the detail is the difference between the sharp Pan image and a smooth intensity component generated as a combination of the interpolated MS bands. In the spatial approach, the detail is the difference between the original Pan image and its version smoothed by a proper lowpass filter, retaining the same spatial frequency content of the MS bands. The dual classes of spectral and spatial methods exhibit complementary features in terms of tolerance to spatial and spectral impairments, respectively [10,11].

The Pan image is preliminarily histogram-matched, that is, radiometrically transformed by a constant gain and offset, in such a way that its lowpass version exhibits mean and variance equal to those of the component that shall be replaced [12]. The injection model rules the combination of the lowpass MS image with the spatial detail of Pan. Such a model is stated between each of the resampled MS bands and a lowpass version of the Pan image having the same spatial frequency content as the MS bands; a contextual adaptivity is generally beneficial [13]. The multiplicative, or contrast-based, injection model with haze correction [14,15] the key to improving the fusion performance by exploiting the imaging mechanism through atmosphere [16]. The injection model, which can rely on the most disparate criteria [17], is crucial for multimodal fusion, where the enhancing and enhanced datasets are produced by different physical imaging mechanisms, such as thermal sharpening [18]. The basic classification of CS and MRA has been progressively

upgraded by considering several other methods that have been recently developed [6], such as those based on Bayesian inference [19], total variation (TV) regularization [20] and sparse representations [21]. More recently, machine learning paradigms have been introduced: since the pioneering study on pansharpening based on convolutional neural networks (CNN) [22], up to extremely sophisticated architectures, such as generative adversarial networks (GAN) [23]. It is noteworthy that, at least for methods based on learning concepts, histogram matching and detail-injection modeling are learned from the training data and implicitly performed by the network, without any control from the user. GAN architectures, however, are able to control one another, and thus, they are invaluable, e.g., for multimodal fusion [24].

This work deals with the use of nonlinear intensity components in spectral MS pansharpening methods. In fact, while CS methods, whose intensity components are linear combinations of the input bands, have been extensively investigated [1], nonlinear intensities have been seldom considered in the literature [25–27]. The hyperspherical color space (HCS) fusion technique [25] is perhaps the most widely known example. Analogously to the linear case, we propose a multivariate linear regression between the interpolated MS bands and the lowpass-filtered Pan. This time, however, the MS and Pan values are squared, before the regression is calculated. Hence, the intensity component no longer lies on a hyperplane in the vector space of the MS samples, as for the linear case, but on a hyper-ellipsoid. The proposed nonlinear intensity component is used in conjunction with the multiplicative injection model. Hence, the de-hazing procedure is extended to the nonlinear intensity.

In an experimental setup, GeoEye-1 and WorldView-3 data in spectral radiance format are pansharpened by several state-of-the-art and up-to-date methods, whose implementations are available in the Pansharpening Toolbox, originally conceived in [28]. The proposed method outperforms all the benchmarks on both the datasets, for all quality indexes at reduced resolution; in particular, its counterpart with linear intensity, referred to as Brovey transform with haze correction (BT-H) [15].

The remainder of this article is organized as follows. Section 2 provides the essential basics of pansharpening. Section 3 introduces the nonlinear intensity and describes the novel method. Section 4 is devoted to haze estimation of MS bands. Section 5 summarizes the adopted quality criteria and the related distortion indexes. Section 6 describes the two datasets and reports simulations and comparisons. Concluding remarks are presented in Section 7.

2. Basics of CS and Mra Pansharpening

Classical pansharpening methods can be divided into CS, MRA and hybrid methods. The unique difference between CS and MRA is the way to extract the Pan details, either by processing the stack of bands along the *spectral* direction or in the *spatial* domain. Hybrid methods, e.g., [29], are the cascade of CS and MRA, either CS followed by MRA or, more seldom, MRA followed by CS. In the former case, they are equivalent to MRA that inherits the injection model from the CS; in the latter case, they behave like CS, with the injection model borrowed from MRA [12]. Therefore, hybrid methods are not a third class with specific properties. The notation used in this paper will be firstly shown. Afterwards, a brief review of CS and MRA will follow.

2.1. Notation

The math notation used is detailed in the following. Vectors are indicated in bold lowercase (e.g., \mathbf{x}) with the i th element indicated as x_i . 2-D and 3-D arrays are denoted in bold uppercase (e.g., \mathbf{X}). An MS image $\mathbf{M} = \{\mathbf{M}_k\}_{k=1,\dots,N}$ is a 3-D array composed of N spectral bands indexed by the subscript k . Accordingly, \mathbf{M}_k indicates the k th spectral band of \mathbf{M} . The Pan image is a 2-D matrix indicated as \mathbf{P} . The MS interpolated and pansharpened MS bands are denoted as $\widetilde{\mathbf{M}}_k$ and $\widehat{\mathbf{M}}_k$, respectively. Unlike conventional matrix product

and ratio, such operations are intended as product and ratio of terms of the same positions within the array.

2.2. CS

The class of CS, or *spectral*, methods is based on the projection of the MS image into another vector space, by assuming that the forward transformation splits the spatial structure and the spectral diversity into separate components.

Under the hypothesis of substitution of a single component that is a linear combination of the input bands, the fusion process can be obtained without the explicit calculation of the forward and backward transformations, but through a proper injection scheme, thereby leading to the fast implementation of CS methods, whose general formulation is [1]:

$$\widehat{\mathbf{M}}_k = \widetilde{\mathbf{M}}_k + \mathbf{G}_k \cdot (\overline{\mathbf{P}}^{(I_L)} - \mathbf{I}_L), \quad k = 1, \dots, N \quad (1)$$

in which k is the band index, $\mathbf{G} = [\mathbf{G}_1, \dots, \mathbf{G}_k, \dots, \mathbf{G}_N]$ the 3-D array of injection gains, which in principle may be different for each pixel each band, while the intensity, \mathbf{I}_L , is defined as

$$\mathbf{I}_L = \sum_{i=1}^N w_i \cdot \widetilde{\mathbf{M}}_i \quad (2)$$

in which the weight vector $\mathbf{w} = [w_1, \dots, w_i, \dots, w_N]$ is the 1-D array of spectral weights, corresponding to the first row of the forward transformation matrix [1]. The term $\overline{\mathbf{P}}^{(I_L)}$ is \mathbf{P} histogram-matched to \mathbf{I}_L

$$\overline{\mathbf{P}}^{(I_L)} \triangleq (\mathbf{P} - \mu_{\mathbf{P}}) \cdot \frac{\sigma_{\mathbf{I}_L}}{\sigma_{\mathbf{P}_L}} + \mu_{\mathbf{I}_L} \quad (3)$$

in which μ and σ denote the mean and square root of variance, respectively, and \mathbf{P}_L is a lowpass version of \mathbf{P} having the same spatial frequency content as \mathbf{I}_L [12].

The simplest CS fusion method is the Intensity–Hue–Saturation (IHS) [1], or better, its generalization to an arbitrary number of bands, GIHS, which allows a fast implementation, given by Equation (1) with unitary injection gains, $\mathbf{G}_k = 1$.

The multiplicative or *contrast-based* injection model is a special case of Equation (1), in which space-varying injection gains, \mathbf{G} , are defined such that

$$\mathbf{G}_k = \frac{\widetilde{\mathbf{M}}_k}{\mathbf{I}_L}, \quad k = 1, \dots, N. \quad (4)$$

The resulting pansharpening method is described by

$$\widehat{\mathbf{M}}_k = \widetilde{\mathbf{M}}_k + \frac{\widetilde{\mathbf{M}}_k}{\mathbf{I}_L} \cdot (\overline{\mathbf{P}}^{(I_L)} - \mathbf{I}_L) = \widetilde{\mathbf{M}}_k \cdot \frac{\overline{\mathbf{P}}^{(I_L)}}{\mathbf{I}_L}, \quad k = 1, \dots, N \quad (5)$$

In Gram–Schmidt (GS) [30] spectral sharpening, the fusion process is described by Equation (1), where the injection gains are spatially uniform for each band, and thus denoted as $\{g_k\}_{k=1, \dots, N}$. They are given by [12]:

$$g_k = \frac{\text{cov}(\widetilde{\mathbf{M}}_k, \mathbf{I}_L)}{\text{var}(\mathbf{I}_L)} \quad k = 1, \dots, N \quad (6)$$

in which $\text{cov}(\mathbf{X}, \mathbf{Y})$ indicates the covariance between \mathbf{X} and \mathbf{Y} , and $\text{var}(\mathbf{Y})$ is the variance of \mathbf{Y} . A multivariate linear regression has been exploited to model the relationship between the lowpass-filtered Pan, \mathbf{P}_L , and the interpolated MS bands [1,31].

$$\mathbf{P}_L = \sum_{i=1}^N \widehat{w}_i \cdot \widetilde{\mathbf{M}}_i + \widehat{b} + \epsilon \triangleq \widehat{\mathbf{I}}_L + \epsilon \quad (7)$$

in which \hat{I}_L is the optimal intensity component and ϵ is the least squares (LS) space-varying residue. The set of space-constant optimal weights, $\{\hat{w}_k\}_{k=1,\dots,N}$, and \hat{b} , is calculated as the *minimum MSE* (MMSE) solution of Equation (7). A figure of merit of the matching achieved by Equation (7) is given by the coefficient of determination (CD), namely R^2 , defined as

$$R^2 \triangleq 1 - \frac{\sigma_\epsilon^2}{\sigma_{\mathbf{P}_L}^2} \quad (8)$$

in which σ_ϵ^2 and $\sigma_{\mathbf{P}_L}^2$ denote the variance of the (zero-mean) LS residue, ϵ , and of the lowpass-filtered Pan image, respectively. Histogram-matching of Pan to the MMSE intensity component, \hat{I}_L , should take into account that $\mu_{\mathbf{P}} = \mu_{\mathbf{P}_L} = \mu_{\hat{I}_L}$, from Equation (7). Thus, from the definition of CD in Equation (8)

$$\bar{\mathbf{P}}^{(\hat{I}_L)} = (\mathbf{P} - \mu_{\mathbf{P}}) \cdot R + \mu_{\mathbf{P}}. \quad (9)$$

This brief review does not include a popular CS fusion method employing principal component analysis (PCA). The reason is that PCA is a particular case of the more general GS transformation, in which I_L is equal to the maximum-variance first principal component, PC1, and the injection gains are those of GS (Equation (6)) [1].

2.3. MRA

The *spatial* approach relies on the injection of highpass spatial details of Pan into the resampled MS bands. The most general MRA-based fusion may be stated as:

$$\hat{\mathbf{M}}_k = \tilde{\mathbf{M}}_k + \mathbf{G}_k \cdot \left(\bar{\mathbf{P}}^{(\tilde{\mathbf{M}}_k)} - \bar{\mathbf{P}}_L^{(\tilde{\mathbf{M}}_k)} \right), \quad k = 1, \dots, N. \quad (10)$$

in which the Pan image is preliminarily histogram-matched to the interpolated k th MS band [12]

$$\bar{\mathbf{P}}^{(\tilde{\mathbf{M}}_k)} \triangleq (\mathbf{P} - \mu_{\mathbf{P}}) \cdot \frac{\sigma_{\tilde{\mathbf{M}}_k}}{\sigma_{\mathbf{P}_L}} + \mu_{\tilde{\mathbf{M}}_k} \quad (11)$$

and $\bar{\mathbf{P}}_L^{(\tilde{\mathbf{M}}_k)}$ the lowpass-filtered version of $\bar{\mathbf{P}}^{(\tilde{\mathbf{M}}_k)}$. It is noteworthy that according to either Equation (3) or Equation (11), histogram matching of \mathbf{P} always implies the calculation of its lowpass version \mathbf{P}_L .

According to Equation (10) the different approaches and methods belonging to this class are uniquely characterized by the lowpass filter employed for obtaining the image \mathbf{P}_L , by the presence or absence of a decimator/interpolator pair [32] and by the set of space-varying injection gains, either spatially uniform, $\{g_k\}_{k=1,\dots,N}$ or space-varying, $\{\mathbf{G}_k\}_{k=1,\dots,N}$.

The contrast-based version of MRA pansharping is

$$\begin{aligned} \hat{\mathbf{M}}_k &= \tilde{\mathbf{M}}_k + \frac{\tilde{\mathbf{M}}_k}{\bar{\mathbf{P}}_L^{(\tilde{\mathbf{M}}_k)}} \cdot \left(\bar{\mathbf{P}}^{(\tilde{\mathbf{M}}_k)} - \bar{\mathbf{P}}_L^{(\tilde{\mathbf{M}}_k)} \right) \\ &= \tilde{\mathbf{M}}_k \cdot \frac{\bar{\mathbf{P}}^{(\tilde{\mathbf{M}}_k)}}{\bar{\mathbf{P}}_L^{(\tilde{\mathbf{M}}_k)}}, \quad k = 1, \dots, N. \end{aligned} \quad (12)$$

It is noteworthy that, unlike what happens for Equation (5), Equation (12) does not preserve the spectral angle of $\tilde{\mathbf{M}}_k$ because the multiplicative sharpening term depends on k , through Equation (11).

Eventually, the projective injection gains derived from GS in Equation (6) can be extended to MRA fusion as

$$g_k = \frac{\text{cov}(\widetilde{\mathbf{M}}_k, \mathbf{P}_L)}{\text{var}(\mathbf{P}_L)} \quad k = 1, \dots, N \tag{13}$$

whose space varying-version \mathbf{G}_k , with statistics calculated locally on a sliding window, coupled with a pyramid MRA constitutes a popular pansharpening method known as GLP-CBD [1,28].

Eventually, we wish to remind that spatial methods are favored if the spatial low-pass filters are designed in order to match the shape of the modulation transfer function (MTF), which is the spatial response in the Fourier domain of the imaging instrument and determines the amount of spatial information conveyed by each spectral channel [33].

3. Pansharpening Based on Nonlinear Intensity Components

The critical review of the baseline HCS [25] is based on the subsequent study by Tu et al. [26], who highlighted the advantages and limitations of the HCS approach. The first idea was to use the HCS transformation as an alternative to the intensity-hue-saturation (IHS) transformation, which had been already generalized to an arbitrary number of bands [1,26], as GIHS. Unfortunately, IHS features a unitary detail-injection model, which is generally poorer than the projection model (GS) and the multiplicative model of Brovey transform (BT) [1,34]. Therefore, in the subsequent publication [26], a fast multiplicative version of HCS was proposed. Fast because it is no longer necessary to calculate the direct and inverse hyperspherical transforms, but only the radius, which is used as intensity component of BT.

The (fast) HCS fusion [26] is given by Equation (5), with the MMSE intensity \mathbf{I}_L replaced by the HCS intensity, $\mathbf{I}_L^{\text{HCS}}$, the radius of the N -dimensional hypersphere:

$$\mathbf{I}_L^{\text{HCS}} = \sqrt{\sum_{i=1}^N \widetilde{\mathbf{M}}_i^2} \tag{14}$$

The original contribution of the present study is to generalize the multivariate regression of Equation (7) to the case of Euclidean distance, as in Equation (14). The result is a new nonlinear intensity component, given by the RMS weighted value of the interpolated MS bands:

$$\widehat{\mathbf{I}}_L^{\text{HECS}} = \sqrt{\sum_{i=1}^N \widehat{w}_i \widetilde{\mathbf{M}}_i^2 + \widehat{b}} \tag{15}$$

in which the set of N spectral weights, \widehat{w}_k , and the bias \widehat{b} are found as the LS solution of the linear regression between squared MS and squared lowpass-filtered Pan:

$$\mathbf{P}_L^2 = \sum_{i=1}^N \widehat{w}_i \cdot \widetilde{\mathbf{M}}_i^2 + \widehat{b} + \boldsymbol{\epsilon} \triangleq \left(\widehat{\mathbf{I}}_L^{\text{HECS}}\right)^2 + \boldsymbol{\epsilon} \tag{16}$$

In the case of three bands, the color spaces of contrast-based fusion methods with linear and nonlinear intensities calculated with and without regression are displayed in Figure 1a–d. Notice that the linear intensity with prefixed equal weights of BT defines an equilateral triangle, as the intersection of a plane with the the first octant of the Euclidean space; the linear intensity with LS weights of BT-H [15] generally yields a scalene triangle. Conversely, as the color space of HCS is the section of a spherical surface lying in the first octant, the proposed method yields an ellipsoidal section. In fact, Equation (15) defines a hyper-ellipsoid, a generalization of the hypersphere in Figure 1c, when the weights may no longer be equal to one another; hence the name hyper-ellipsoidal color space (HECS) fusion.

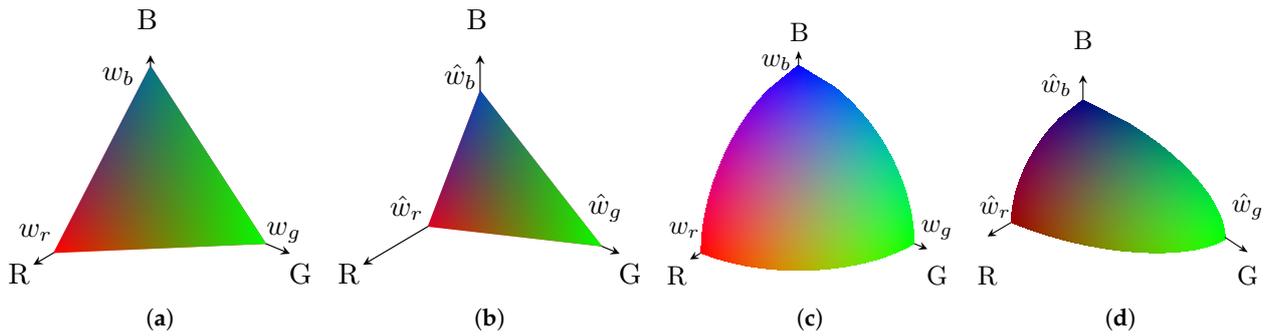


Figure 1. Linear and nonlinear intensity components in a three-band color space: (a) BT; (b) BT-H; (c) HCS; (d) HECS.

The proposed scheme includes de-hazing, highly beneficial for fusion methods with a multiplicative detail-injection model [15,35]. Hence, the formulation of the proposed HECS pansharpening fusion is

$$\widehat{\mathbf{M}}_k = \left(\widetilde{\mathbf{M}}_k - L_p(k) \right) \cdot \frac{\widehat{\mathbf{P}}^{(\widehat{\mathbf{I}}_L^{\text{HECS}})} - L_p(\widehat{\mathbf{I}}_L^{\text{HECS}})}{\widehat{\mathbf{I}}_L^{\text{HECS}} - L_p(\widehat{\mathbf{I}}_L^{\text{HECS}})} + L_p(k), \quad k = 1, \dots, N \quad (17)$$

in which the path-radiance, or *haze*, of the synthetic intensity, $L_p(\widehat{\mathbf{I}}_L^{\text{HECS}})$, is given by the weighted RMS value of the individual path radiances, $L_p(k), k = 1, \dots, N$,

$$L_p(\mathbf{P}) = L_p(\mathbf{P}_L) = L_p(\widehat{\mathbf{I}}_L^{\text{HECS}}) = \sqrt{\sum_{i=1}^N \widehat{w}_i \cdot L_p^2(i) + \widehat{b}}. \quad (18)$$

Equation (17) shows that the spectral pixel vector is translated by the haze vector before the multiplicative fusion is accomplished and the fused pixel vector is translated back by the same haze vector. Figure 2 shows a flowchart describing the HECS fusion process. The estimation of the atmospheric path radiances of the individual MS bands will be tackled in the next section.

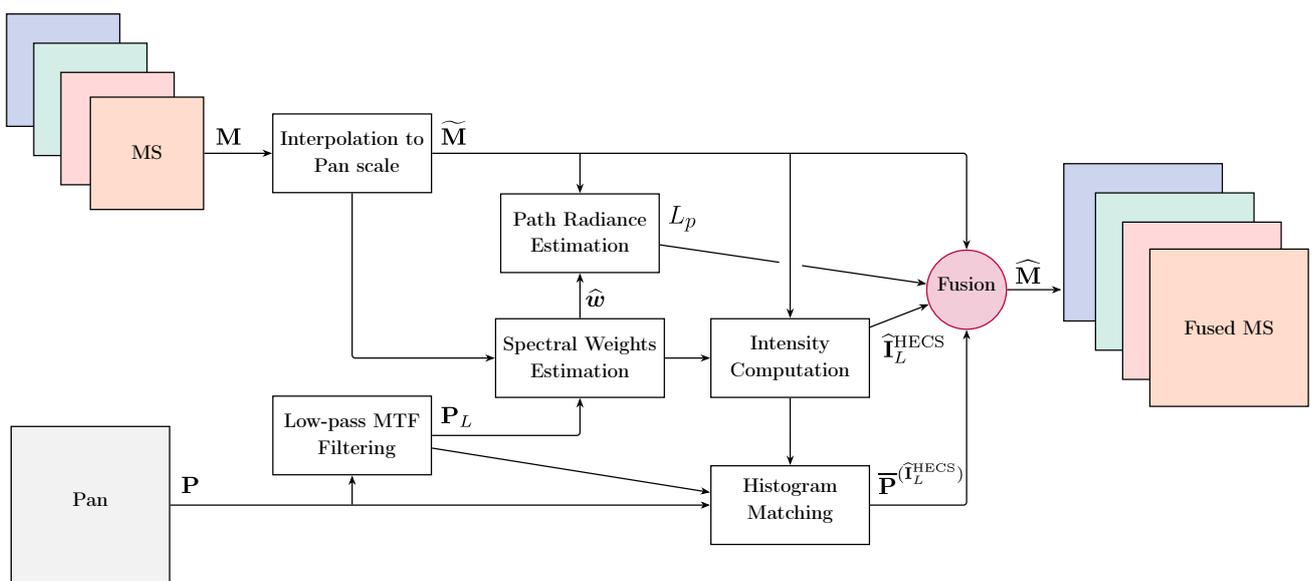


Figure 2. Flowchart of the proposed HECS pansharpening.

4. Haze Estimation

MS pansharpening, which produces a sharp MS image having the same format as the original MS image [12], generally does not require any kind of atmospheric corrections, unless a multiplicative detail-injection model is adopted [15]. In this case, it was proven that the haze-corrected pansharpening is capable of thoroughly preserving the map of normalized differential vegetation index (NDVI) of the original MS data [35].

4.1. Definition of Haze

All the different atmospheric constituents, as natural or anthropogenic aerosols that also have other important effects on population [36] and climate changes [37], gases (e.g., nitrogen, oxygen), and hydro-meteors (e.g., liquid and ice clouds, precipitation) scatter and absorb the incoming shortwave solar radiation. Haze is then the responsible for all the unwanted wavelength-dependent backscattered solar radiation by the previously described atmospheric components that does not go under an interaction with the Earth's surface. Haze term correction requires path-radiance value estimation/calculation for the different spectral bands. Path-radiance estimation may follow image-based approaches or rely on radiative transfer models of the atmosphere and its constituents, as well as on knowledge of acquisition parameters, such as the actual Sun–Earth distance, solar zenith angle and satellite platform observation angle.

Image-based atmospheric correction techniques [38] are a series of statistical methods based on some general assumptions and empirical criteria. The goal is estimating the atmospheric effects on acquisition without requiring acquisition parameters or making assumptions on atmospheric constituents. If the band offsets (metadata) in the file header are all identically zero, image-based methods are also directly applicable without a preliminary conversion. In this case, the path-radiance values estimated for each band will not be expressed in physical units, but as DN_s. Conversion to spectral radiance units, typically [$W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$] or [$mW \cdot cm^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$] requires subsequent multiplication by the calibration gain metadata. If the offsets are nonzero, it is preferable to convert all pixel values to physical units before applying image-based methods.

4.2. Shadowed Pixel Assumption

In order to apply image-based methods to estimate the path-radiance values, let us make a series of assumptions: (i) the path radiance is uniform over the scene; (ii) the scene is large enough, in terms of number of pixels, and hence statistically consistent; (iii) the scene contains shadowed pixels where the direct solar irradiance is masked by some obstacle; (iv) the diffuse irradiance at the surface is an order of magnitude smaller with respect to the direct solar irradiance. Under those assumptions, the path radiance of the k th band will be equal to the minimum of spectral radiance attained over the k th band of the scene. In practice, instead of the minimum (0-percentile) the haze is taken equal to the 1-percentile of the histogram. This strategy is robust and permits us to deal with the small, but non-zero, diffuse irradiance and to the photon and thermal instrument noise appearing as fluctuations of the dark signal around its average, and to outliers originated by pattern-gain correction of the instrument. The spatial scale of representation is crucial: a full shadowed pixel at higher spatial resolution might not be anymore when mapped onto larger pixels. The path radiance arguably does not depend on the spatial scale. Scale invariance at 2 m and 8 m is attained considering percentile values between 0.5 and 1. Below 0.5, the invariance is weak while above 1, the invariance is almost perfect. Thus, the path radiance of the k th band, which is usually approximated by its minimum (0-percentile) over the scene [38,39], may be better approximated by any value between the 0.5-percentile and the 1-percentile.

If the scene does not contain at least one pure shadow pixel, i.e., snow-covered surfaces or desert environment, the above method may lead to serious estimation errors. A shortcoming would be assuming that at least one pixel in the scene has a reflectance close to zero [39]. This assumption is unlikely for all wavelengths, but may approximately hold

for the blue (B) band, where the atmospheric Rayleigh scattering is ten times stronger with respect to the red band (R).

4.3. Haze Computation

So, once the path-radiance of the B channel, $L_p(B)$, is known, being estimated as a small percentile of the local histogram, the intercept of the scatterplot of G versus $(B - L_p(B))$ is an estimate of the path radiance of the green (G) channel, $L_p(G)$; analogously for the red (R) channel, $L_p(R)$ may be estimated from the R-to- $(G - L_p(G))$ scatterplot. The scatterplot method holds for the visible bands [38]. For calculating the path radiance of NIR, which is practically uncorrelated with the visible bands in the presence of vegetation [40], the scatterplot method may fail, unless its calculation, either supervised or unsupervised, e.g., NDVI-enforced, is performed on non-vegetated areas. Otherwise, a reasonable physical approximation is that the $L_p(\text{NIR})$ is set equal to zero.

Additionally, a modeling of the atmosphere is an option. In this case, the DNs must be preliminarily calibrated by means of the gains and offsets metadata. Radiative transfer modeling requires acquisition year, month, day, local time, longitude, latitude, and possibly type of landscape for setting aerosols (advected [41] or local) both in the boundary layer or upper troposphere. The content of water vapor may be inferred from estimation of the rain evaporation rate [42]. The Fu–Liou–Gu (FLG) model [43] directly yields values of path-radiance in predefined bands, roughly corresponding to those of MS scanners, such as Landsat 8 OLI. With small adjustments to fit the R and NIR bands of GeoEye-1, it was recently found [35] that the modeled path radiance is well approximated by 95% of the 1-percentile of B, 65% of the 1-percentile of G, 45% of the 1-percentile of R and 5% of the 1-percentile of NIR.

An exhaustive search at steps of one DN was performed in [15]. The optimal values of path-radiances are those that optimize fusion scores, on average, at the degraded spatial scale, i.e., when the ground truth is available as reference. With FLG-modeled path radiances and model-enforced image-based path-radiances, the performance is about 0.1 % lower than that achieved with the exhaustive search. Therefore, the accuracy of path radiance estimation is not crucial, at least for clear atmospheres. Interestingly, the fusion performance plots vs. corrected path-radiance values exhibit a maximum that is peaked for under-corrected haze and much flatter for over-corrected haze. This fact entails the use of band minima as estimated haze values, that is, $L_p(k) = \min(\mathbf{M}_k), k = 1, \dots, N$.

Trivially, the haze term is zero for data in surface reflectance format, whenever they are available. The surface reflectance is a level-two (L2) product and is distributed for global-coverage systems (OLI, Sentinel-2), only where an instrument network is available for atmospheric measurements [44], usually carried out by means of lidar instruments [45,46].

5. Quality Assessment

Quality evaluation of image fusion products has been, and still is, the object of extensive research. The problem is complicated by the fact that it may not be easy to formalize what “quality” means in the fusion process. In this regard, a protocol of assessment should have very clear objectives and possibly require a reference on which the comparison relies. Image fusion assessment is traditionally performed in two ways: (1) by means of a human visual inspection by a panel of investigators; and (2) through mathematical functions capable of objectively measuring or inferring the similarity of the fusion product to a reference target, which is always unavailable and often also undefined. Whereas the former is based on subjective human evaluations that can be embodied by some statistical indexes, such as entropy, contrast, gradient, and so on [24], the objective evaluation involves stringent and quantitative measures that involve both original and fused images and are possibly consistent with human visual perception. In the fusion of medical images, it is crucial to preserve the diagnostic characteristics of the original images within the fused image; thus, it is necessary to evaluate the result of the fusion process using objective parameters.

In remote sensing image fusion, especially multispectral pansharpening applications, quality assessment is performed following Wald's protocol [47], which substantially requires the fused image to satisfy three main properties:

- *Consistency*: the fused image, once spatially degraded to the original resolution, should be as close as possible to the original image;
- *Synthesis*: any low-resolution (LR) image fused by means of a high-resolution (HR) image should be as identical as possible to the ideal image that the corresponding sensor, if existent, would observe at the resolution of the HR image.
- *Vector synthesis*: the set of multispectral images fused by means of the HR image should be as identical as possible to the set of ideal images that the corresponding sensor, if existent, would observe at the spatial resolution of the HR image.

The property of consistency is usually easier to assess since the original LR image can be used as a reference. Only the procedure of spatial degradation and the matching function are to be standardized. On the contrary, the synthesis property is harder to be verified, since a reference is required. A viable shortcoming stems from the assumption of scale-invariance of the scene, that is, quality measures do not vary with the resolution, at which the scene is imaged. This allows the quality to be measured at a resolution lower than the original one, for which the reference image is available.

More specifically, the process consists of spatially degrading both the enhancing and the enhanced datasets by a factor equal to the scale ratio between them and using the original LR image as reference. Obviously, such an assumption is not always valid, especially when the degradation process does not mimic the actual sensor acquisition process. In the case of multimodal image fusion, the applicability of the synthesis properties of the Wald's protocol is questionable since a multimodal fusion method aims at producing images in which the features coming from different sensors should in principle be both present. If the imaging sensors exploit different physical mechanisms, e.g., reflectivity and emissivity in the case of fusion of optical and thermal data, the assumption that an "ideal" sensor producing the fused image could exist is unlikely, since such a sensor should be able to measure and integrate different physical phenomena at the same time.

In conclusion, notwithstanding achievements over years [48–54], quality assessment of pansharpened images is still an open problem, being inherently ill-posed. A further source of uncertainty, which has been explicitly addressed very seldom [55,56], is that the measured quality may also depend on the data format.

5.1. Reduced-Resolution Assessment

The quality check often entails the shortcoming of performing fusion with both MS and Pan datasets degraded at spatial resolutions lower than those of the originals, in order to use non-degraded MS originals as quality references [57]. Here, some popular statistical similarity/dissimilarity indexes used in this study will be briefly reviewed.

5.1.1. SAM

The spectral angle mapper (SAM) was originally introduced for the discrimination of materials starting from their reflectance spectra [58]. Given two spectral vectors, \mathbf{v} and $\hat{\mathbf{v}}$, both having N components, in which $\mathbf{v} = \{v_1, v_2, \dots, v_N\}$ is the reference spectral pixel vector and $\hat{\mathbf{v}} = \{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_N\}$ is the test spectral pixel vector, SAM denotes the absolute value of the spectral angle between the two vectors:

$$\text{SAM}(\mathbf{v}, \hat{\mathbf{v}}) \triangleq \arccos\left(\frac{\langle \mathbf{v}, \hat{\mathbf{v}} \rangle}{\|\mathbf{v}\|_2 \cdot \|\hat{\mathbf{v}}\|_2}\right). \quad (19)$$

SAM is usually expressed in degrees and is equal to zero if the test vector is spectrally identical to the reference vector, i.e., the two vectors are parallel and may differ only by their moduli. A global spectral dissimilarity, or distortion, index is obtained by averaging Equation (19) over the scene.

5.1.2. ERGAS

ERGAS, the French acronym for relative dimensionless global error in synthesis [59], is the cumulative normalized root mean square error (NRMSE) between test and reference band, multiplied by the Pan-to-MS scale ratio and expressed in percentage:

$$\text{ERGAS} \triangleq 100 \frac{d_h}{d_l} \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{\text{RMSE}(k)}{\mu(k)} \right)^2} \quad (20)$$

where d_h/d_l is the ratio between pixel sizes of Pan and MS, e.g., 1/4, $\mu(k)$ is the mean (average) of the k th band of the reference and N is the number of bands. Low values of ERGAS indicate high similarity between fused and reference MS data.

5.1.3. Multivariate UIQI

$Q2^n$ is the multiband extension of the universal image quality index (UIQI) [60] and was introduced for quality assessment of pansharpened MS images [61]. Each pixel of an image with N spectral bands is accommodated into a hypercomplex (HC) number with one real part and $N - 1$ imaginary parts.

Let $z = z(m, n)$ and $\hat{z} = \hat{z}(m, n)$ denote the HC representation of the reference and test spectral vectors at pixel (m, n) . Analogously to UIQI, namely, $Q2^0 = Q$, $Q2^n$ may be written as product of three terms:

$$Q2^n = \frac{|\sigma_{z\hat{z}}|}{\sigma_z \sigma_{\hat{z}}} \cdot \frac{2\sigma_z \sigma_{\hat{z}}}{\sigma_z^2 + \sigma_{\hat{z}}^2} \cdot \frac{2|\bar{z}||\bar{\hat{z}}|}{|\bar{z}|^2 + |\bar{\hat{z}}|^2} \quad (21)$$

the first of which is the modulus of the HC correlation coefficient (HCCC) between z and \hat{z} . The second and third terms, respectively, measure contrast changes and mean bias on all bands simultaneously. Statistics are calculated on square blocks, typically, 32×32 , and $Q2^n$ is averaged over the blocks of the whole image to yield the *global* score index. $Q2^n$ takes values in $[0, 1]$ and is equal to 1 iff $z = \hat{z}$ for all pixels.

5.2. Full-Resolution Assessment

Quality can be evaluated at the original panchromatic scale, according to a *full resolution* (FR) approach [62]. In this case, the spectral and spatial distortions are separately evaluated starting from the fused image and either the original low-resolution MS bands or the high-resolution panchromatic image, as firstly proposed by Zhu et al. [63].

5.2.1. QNR

A widely adopted FR assessment is based on the quality with no reference (QNR) protocol [51] and the related distortion indexes. QNR combines into a unique overall quality index a spectral distortion measure between the original and pansharpened MS bands and a spatial distortion measure between each MS band and PAN. The QNR protocol is based on the following assumptions:

1. The fusion process should not change the intra-relationships between couples of MS bands; in other words, any intra-relationship changes between couples of MS bands across resolution scales are considered as indicators of spectral distortions;
2. The fusion process should not change the inter-relationships between each MS band and the Pan image; in other words, any inter-relationship changes between each MS band and the Pan across resolution scales are modeled as spatial distortions.

The QNR protocol employs the UIQI as a similarity measure and the absolute difference as the change operator. The spectral distortion index, D_λ , is obtained by computing

two sets of UIQI values, each between couples of MS bands. Afterward, their absolute difference is taken and averaged:

$$D_\lambda = \frac{1}{N(N-1)} \sum_{l=1}^N \sum_{\substack{r=1 \\ r \neq l}}^N |Q(\widetilde{\mathbf{M}}_l, \widetilde{\mathbf{M}}_r) - Q(\widehat{\mathbf{M}}_l, \widehat{\mathbf{M}}_r)|. \quad (22)$$

The spatial distortion index, D_s , is computed by means of the average absolute UIQI band by band difference, between MS and Pan, both at FR and at the original MS resolution:

$$D_s = \frac{1}{N} \sum_{i=1}^N |Q(\widetilde{\mathbf{M}}_i, \mathbf{P}_L) - Q(\widehat{\mathbf{M}}_i, \mathbf{P})|. \quad (23)$$

Finally, a unique quality index is obtained combining the complement of spatial and spectral distortion indexes:

$$\text{QNR} = (1 - D_\lambda)^\alpha \cdot (1 - D_s)^\beta. \quad (24)$$

The exponents α and β rule the balance of the spectral and spatial quality components. They can be normalized in such a way that $\alpha + \beta = 1$. In this case, if $\alpha = \beta = 0.5$, Equation (24) yields the geometric mean of the spectral and spatial qualities, though normalization of exponents compresses the variability of the cumulative index. Typical values for the exponents are $\alpha = \beta = 1$.

5.2.2. Khan's QNR

A totally different approach was later proposed by Khan et al. [52]. Analogously to QNR, Khan's QNR (KQNR) defines and combines spectral and spatial consistency factors. The innovations introduced by the KQNR protocol is to make use of the consistency property of Wald's protocol to calculate the spectral consistency of the pansharpened product. Since the consistency property evaluation requires a spatial degradation stage, including a decimation operation, the KQNR protocol proposes to use MTF-matched filters to perform the spatial degradation of the fused MS bands. Thus, the spectral distortion index, $D_\lambda^{(K)}$, is computed according to the following procedure:

1. Each fused MS band is spatially degraded (filtered and decimated) with its specific MTF-matched filter;
2. The $Q2^n$ index between the set of spatially degraded fused MS images and the original MS dataset is computed;
3. The one's complement is taken to obtain a distortion measure:

$$D_\lambda^{(K)} = 1 - Q2^n(\widehat{\mathbf{M}}_{L\downarrow}, \mathbf{M}). \quad (25)$$

The spatial consistency of Khan's protocol is given by the average change in interscale similarities between highpass components of each fused band and Pan:

$$D_s^{(K)} \triangleq \frac{1}{N} \sum_{k=1}^N |Q(\widehat{\mathbf{M}}_{kH}, \mathbf{P}_H) - Q(\mathbf{M}_{kH}, (\mathbf{P}_{L\downarrow})_H)|. \quad (26)$$

Again, a cumulative quality index is obtained combining the complement of spatial and spectral distortion indexes:

$$\text{QNR} = (1 - D_\lambda^{(K)})^\alpha \cdot (1 - D_s^{(K)})^\beta \quad (27)$$

with typical values for the exponents $\alpha = \beta = 1$.

It is noteworthy that, unlike what happened with QNR, the KQNR protocol states the the spectral and spatial consistencies are calculated on the lowpass and highpass spatial-frequency channels of the fused images; in the former case with a comparison with the

original MS; in the latter case with a comparison with the highpass components of the original Pan and of the spatially degraded Pan.

5.2.3. Hybrid QNR

The hybrid QNR (HQNR) has been presented in [53] as the combination of the spectral distortion index of KQNR in Equation (25) with the spatial distortion index of QNR in Equation (23).

Analogously to QNR and KQNR, a unique quality index is obtained combining the one's complements of the spectral and spatial distortions:

$$\text{HQNR} = \left(1 - D_{\lambda}^{(K)}\right)^{\alpha} \cdot \left(1 - D_s\right)^{\beta}. \quad (28)$$

For the sake of completeness, we could consider the dual of HQNR (DQNR), in which the spectral distortion in Equation (22) is coupled with the spatial distortion in Equation (26), defined as:

$$\text{DQNR} = \left(1 - D_{\lambda}\right)^{\alpha} \cdot \left(1 - D_s^{(K)}\right)^{\beta}. \quad (29)$$

6. Experimental Results and Discussion

6.1. Data Sets

Two test images, acquired by two different platforms, GeoEye-1 and WorldView-3, have been used in the simulations. In this section, besides describing the two datasets, which belong to the reference pansharpening dataset, namely PAirMax, described in [64], we present an analysis of the solution of the multivariate regressions (Equations (7) and (16)) for the two datasets.

6.1.1. Trenton Dataset

The GeoEye-1 image has been acquired over the area of Trenton NJ, USA, on 27 September 2019. The on-orbit limit sampling interval at nadir (half the system's resolution) is 1.64 m for MS and 0.41 m for Pan. The spatial sampling interval (SSI) of the resampled geocoded product is 2 m for MS and 0.5 m for Pan, resulting in a scale ratio equal to four. The MS image features four spectral bands: blue (B), green (G), red (R) and near infra-red (NIR). The image size is 512×512 pixels for MS and 2048×2048 for Pan. The radiometric resolution of the DN format is 11 bits; the conversion coefficients to the SR format, extracted from the metadata, are reported in Table 1. It is noteworthy that the offsets are all equal to zero.

Table 1. Gains and offsets for conversion to SR of GeoEye-1—Trenton.

GE-1	Pan	B	G	R	NIR
α_k	0.1779	0.1487	0.1718	0.1619	0.0959
β_k	0	0	0	0	0

6.1.2. Munich Dataset

The WorldView-3 image, acquired on 10 January 2020, portrays the city of Munich, Germany, and surrounding agricultural fields and forested areas. The on-orbit limit sampling interval at nadir is 1.24 m for MS and 0.31 m for Pan. The SSI is 1.32 m for MS and 0.33 m for Pan; hence, the scale ratio is still four. The MS image comprises eight bands: coastal (C), B, G, yellow (Y), R, red edge (RE), NIR1 and NIR2. The covered area is $456,759 \text{ m}^2$, corresponding to MS and Pan image sizes of 512×512 and 2048×2048 pixels, respectively. The fixed-point representation of the data employs 11 bits. Table 2 shows the corresponding radiometric calibration coefficients obtained from the metadata file. Moreover, in this case, the offsets are identically equal to zero.

Table 2. Gains and offsets for conversion to SR of WorldView-3—Munich.

WV-3	Pan	C	B	G	Y	R	RE	NIR1	NIR2
α_k	0.1365	0.3451	0.1900	0.1233	0.1764	0.1010	0.1567	0.0675	0.1164
β_k	0	0	0	0	0	0	0	0	0

6.2. Analysis of the LS Intensity Component

The LS solutions of the multivariate linear regression, both Equations (7) and (16), are now discussed for the two test datasets. Tables 3 and 4 report the MMSE spectral weights of each band and the bias coefficients, if they are included, for both the datasets preliminarily converted to SR format. The regression of squares in Equation (16) produces spectral weights that are far different from the other case. However, it is equivalent to Equation (7) in terms of matching success, for Trenton; slightly less fitting for Munich, supposedly, because the bandwidth of Pan does not include the outermost MS bands of WorldView-3. In both regressions, Equations (7) and (16), the presence or absence of the bias term produces \hat{w}_k s that are somewhat different, especially for Munich, but this has very limited impact on the degree of matching, measured by CD, which changes by 10^{-5} for Munich and even less for Trenton.

Table 3. Regression coefficients of GE-1—Trenton, calculated at reduced resolution.

Trenton	B	G	R	NIR	\hat{b}	$\sum \hat{w}_k$	CD
Lin.	−0.0573	0.5384	0.3072	0.2710	−0.0128	1.0593	0.9916
Lin. w/o bias	−0.0582	0.5393	0.3071	0.2709	−	1.0591	0.9916
Quad.	−0.1539	0.6661	0.3140	0.2296	87.8771	1.0558	0.9913
Quad. w/o bias	−0.0820	0.5802	0.3327	0.2437	−	1.0746	0.9913

Table 4. Regression coefficients of WV-3—Munich, calculated at reduced resolution.

Munich	C	B	G	Y	R	RE	NIR1	NIR2	\hat{b}	$\sum \hat{w}_k$	CD
Lin. w/ bias	0.1897	−0.0119	0.0960	0.4025	0.0522	0.2294	−0.0051	0.1955	−3.6256	1.1482	0.9855
Lin. w/o bias	−0.0121	0.1635	0.0442	0.3949	0.0624	0.2352	0.0354	0.1313	−	1.0548	0.9854
Quad. w/ bias	0.0701	0.0386	0.1140	0.3149	0.1195	0.3202	−0.1904	0.5874	−76.2715	1.3742	0.9794
Quad. w/o bias	−0.0311	0.1496	0.0518	0.3339	0.1175	0.3132	−0.1598	0.5318	−	1.3070	0.9793

6.3. Simulations

Twelve pansharpening algorithms, including the plain interpolated MS image, without injection of details, denoted by EXP, have been selected from the Pansharpening Toolbox [6,28]. The list includes:

- Brovey transform (BT) [1,34];
- Gram–Schmidt (GS) spectral sharpening [1,30];
- Fast fusion with hyperspherical color space (HCS) [26];
- Optimized BT with haze correction (BT-H) [15];
- GS with adaptive intensity (GSA) [1,28];
- The proposed method with hyper-ellipsoidal color space (HECS);
- Fusion method with band-dependent spatial-details (BDS) injection [65];
- Additive wavelet luminance proportional with haze correction (AWLP-H) [31];
- GLP with MTF filters and full-scale detail injection modeling (MTF-GLP-FS) [66];
- Fusion based on sparse representation of spatial details (SR-D) [21];
- Fusion based on total-variation (TV) optimization [20];
- Advanced pansharpening with neural networks and fine tuning (A-PNN-FT) [22].

Table 5 reports the numerical results of the RR simulations performed on the two datasets. There is a significant increment in performance of HECS over its baseline HCS. Comparisons with BDSD, BT-H and AWLP-H reveal that the three haze-corrected methods are practically equivalent in performance. The performance of BT is noteworthy for Trenton, but average for Munich. In fact, the intensity component calculated as an average of the interpolated MS bands attains a high degree of matching with the lowpass-filtered Pan: CD equal to 0.9904, very similar to the CDs of BT-H and HECS in Table 3. The high CD and the small amount of vegetated areas, from which haze correction benefits, in the Trenton image, makes all contrast-based CS methods perform very similar to one another, regardless of their intrinsic structure, more or less developed. Overall, HECS outperforms all the other methods in all quality parameters on the Munich test site. On the Trenton image, the performance of A-PNN-FS in terms of SAM stands out.

Table 5. Fusion comparison at reduced resolution for the WorldView-3—Munich and GeoEye-1—Trenton datasets. GT indicates reference ground truth. Best values boldfaced and second best underlined. All algorithms are run in SR format, except A-PNN-FT, which requires the same DN format used for training, whose results were converted to SR before the assessment.

Dataset	Munich				Trenton			
	Q8	Qavg	SAM	ERGAS	Q4	Qavg	SAM	ERGAS
GT	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000
EXP	0.6311	0.6354	4.7548	10.8511	0.5826	0.5894	6.6167	10.2034
BT	0.8803	0.8703	4.7548	5.5754	0.9000	0.8938	6.6167	5.3655
GS	0.8028	0.8190	4.2535	6.9518	0.8461	0.8513	6.2997	6.6388
HCS	0.8906	0.8633	4.7548	6.1731	0.8969	0.8909	6.6167	5.4681
BT-H	0.9236	<u>0.9298</u>	<u>2.9309</u>	4.2466	0.9025	0.9052	4.9937	<u>4.9978</u>
GSA	0.9204	0.9215	3.2007	4.4250	0.8985	0.8962	6.0420	5.2664
HECS	0.9287	0.9347	2.9078	4.1268	0.9066	0.9091	<u>4.9565</u>	4.9609
BDSD	<u>0.9245</u>	0.9269	3.2388	<u>4.1748</u>	<u>0.9054</u>	<u>0.9065</u>	6.0254	5.1267
AWLP-H	0.9154	0.9135	2.9794	4.3915	0.8928	0.8946	5.2913	5.2182
MTF-GLP-FS	0.9200	0.9210	3.1876	4.4465	0.9030	0.9005	6.0093	5.1501
SR-D	0.8936	0.8991	3.4386	5.3399	0.8915	0.8946	5.4449	5.3810
TV	0.9164	0.9190	3.4225	4.6557	0.7693	0.7711	6.1318	7.7066
A-PNN-FT	0.8747	0.8798	3.6465	5.8899	0.8857	0.8895	4.3841	5.4262

The results of the fusion process at reduced resolution for each of the methods, are shown in Figure 3, for the datasets of Munich. HCS is visually sharp, but colors are somewhat distorted. The three haze-corrected methods, BT-H, AWLP-H and HECS look very similar to one another. BDSD is fine but the tree canopies are unrealistically textured. SR-D surprisingly exhibits striping artifacts in built-up areas.

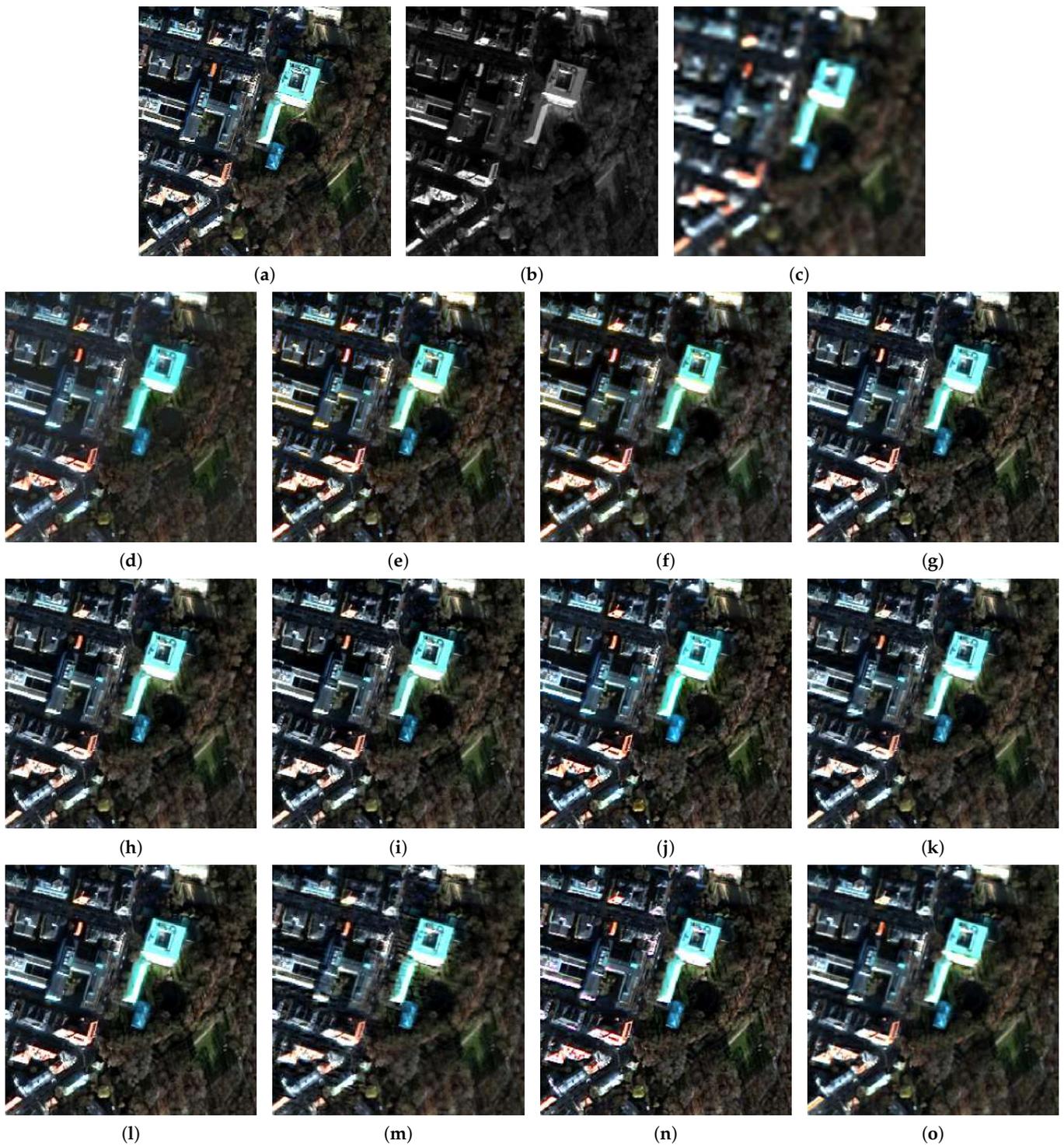


Figure 3. Fusion results at reduced resolution, using a true-color representation, for the Munich dataset: (a) Reference; (b) Pan image; (c) expanded; (d) GS; (e) BT; (f) HCS; (g) GSA; (h) BT-H; (i) HECS; (j) BDSD; (k) AWLP-H; (l) MTF-GLP-FS; (m) SR-D; (n) TV; (o) A-PNN-FT.

Table 6 reports the numerical results of the FR simulations performed on the Trenton dataset. Such results are reported for the four main distortions, two spectral and two spatial, and for all the four possible combinations of a spectral and a spatial distortion, considered in Section 5.2. Apart from EXP, SR-D, TV and A-PNN-FT, which do not follow the strict classification of spectral and spatial methods, BT, GS, HCS, BT-H, GSA, HECS and BDSD are methods based on CS; AWLP-H and MTF-GLP-FS on MRA. In the ranking of methods

provided by QNR, BSDS is the first method, TV the second best and HECS the third one. The ranking provided by KQNR and HQNR is far different: HECS is at the sixth place for KQNR and even at the seventh for HQNR. Eventually, DQNR provides a totally different ranking, more similar to that of QNR, in which HECS is in first position, closely followed by AWLP-H. This puzzling behavior of the FR quality indexes has a twofold explanation: (i) the spatial distortion of QNR in Equation (23) was found to be very sensitive to the choice of the decimation filter of Pan [57], which is unknown in FR experiments; (ii) according to Khan’s protocol, all CS methods exhibit very high values of spectral distortion (Equation (25)) and the sole possible explanation [6,62] is that, due to the presence of hardly perceivable space-varying misalignments between interpolated MS and Pan, CS methods shift the fused MS towards Pan, thereby losing its consistency towards the original MS [8,67]. If we restrict the comparison of HECS to CS methods only, we will find that for all cumulative indexes containing the spectral distortion in Equation (25), namely KQNR and HQNR, HECS is the best among CS methods, though poorer than those methods that are not CS. We can conclude that in quantitative assessments carried out at FR, HECS confirms its superiority, though moderate, over benchmarks that represent the state-of-the-art of MS pansharpening.

Table 6. Fusion results of GE-1—Trenton at full resolution. Best values boldfaced and second best underlined. Boxed boldfaced and boxed underlined denote best and second best values restricted to the category of CS methods. All algorithms are run in SR format, except A-PNN-FT, which requires the same DN format used for training, whose results were converted to SR before the assessment.

	D_λ	D_s	QNR	$D_\lambda^{(K)}$	$D_s^{(K)}$	KQNR	HQNR	DQNR
EXP	0.0000	0.0938	0.9062	0.0887	0.1844	0.7432	0.8258	0.8156
BT	0.0269	0.0924	0.8831	0.1457	0.0472	0.8140	0.7754	0.9272
GS	<u>0.0171</u>	0.0834	0.9009	0.1535	0.0688	0.7882	0.7759	0.9153
HCS	0.0306	0.0851	0.8869	<u>0.1472</u>	0.0418	<u>0.8171</u>	0.7802	0.9289
BT-H	0.0305	0.0830	0.8890	0.1480	0.0451	0.8136	<u>0.7813</u>	0.9258
GSA	0.0456	0.1023	0.8567	0.1533	0.0528	0.8020	0.7600	0.9040
HECS	0.0221	0.0606	0.9187	0.1592	0.0275	0.8177	0.7899	0.9510
BSDS	0.0339	0.0135	0.9530	0.2171	0.0745	0.7246	0.7723	0.8941
AWLP-H	0.0463	0.0468	0.9090	0.0525	<u>0.0106</u>	<u>0.9375</u>	<u>0.9032</u>	<u>0.9436</u>
MTF-GLP-FS	0.0727	0.0651	0.8670	<u>0.0505</u>	0.0102	0.9397	0.8877	0.9178
SR-D	0.0843	0.0816	0.8409	0.0314	0.0656	0.9051	0.8896	0.8556
TV	0.0233	0.0374	<u>0.9402</u>	0.0776	0.1015	0.8288	0.8879	0.8776
A-PNN-FT	0.0774	<u>0.0300</u>	0.8949	0.0629	0.0404	0.8993	0.9090	0.8853

The results of the FR fusion process are shown in Figure 4, for the datasets of Trenton. GS, BT, HCS and GSA are very similar to one another: clean and geometrically fine, but with little distortion of color spots. BT-H and HECS are very similar: geometrically accurate, adequately textured (thanks to dehazing), with spectral brightness of color spots, slightly better for HECS (see colored cars). BSDS is finely textured, with bright color spots, but little realistic. AWLP-H and MTF-GLP-FS are rather fine, but less geometrically accurate than CS methods, presumably due to misalignment of datasets. SR-D is unrealistically over-textured. TV is accurate, but slightly little sharp. A-PNN-FT is generally fine, though the typical glitch effects of neural methods are visible. In conclusion, the visual evaluations substantially agree with the ranking of methods achieved by means of DQNR. The ranking of QNR is partially flawed. KQNR and HQNR exhibit little agreement with visual evaluations, because spectral consistency measurements are impaired by the imperfect alignment.



Figure 4. Fusion results at full resolution, using a true-color representation, for the Trenton dataset: (a) Pan image; (b) expanded; (c) GS; (d) BT; (e) HCS; (f) GSA; (g) BT-H; (h) HECS; (i) BDSD; (j) AWLP-H; (k) MTF-GLP-FS; (l) SR-D; (m) TV; (n) A-PNN-FT.

6.4. Discussion

After a careful analysis of RR and FR results, both numerical indexes and true-color icons, a series of considerations can be made. Among methods featuring the multiplicative detail injection model, those performing haze correction (BT-H, HECS and AWLP-H) provide superior performance in the presence of vegetated areas. On the urban landscape

of Trenton, the difference in performance, as measured in Table 5, among BT, BTH-H, HCS, HECS and AWLP-H is small.

The regression is crucial for the calculation of the injected detail, not for the injection model. In this sense, BT-H is slightly better than AWLP-H, where the two methods have exactly the same injection model and the same haze correction.

The question that arises is: why the proposed HECS is slightly better than BT-H, its counterpart employing a *linear* intensity (see Figure 1b,d). In Equation (17), the narrow-band haze terms are the same as for BT-H, the definition of the intensity component and the broadband haze terms of Equation (18) are different from those of BT-H, but identical at numerator and denominator of the ratio in Equation (17). If the broadband haze terms of HECS and BT-H are swapped, both methods slightly lose performance. Histogram matching of Pan to the intensity component is standard, as in Equation (3).

A reasonable explanation is that the regression of quadratic values in Equation (17) is such that the resulting LS intensity component better matches the lowpass-filtered Pan than the intensity achieved through a linear regression. With reference to Tables 3 and 4, the CDs of the quadratic cases are slightly lower than those of the corresponding linear cases, but in the former case CD, measures the extent to which the *squared* intensity matches the *squared* lowpass-filtered Pan, in the latter case the extent to which the intensity matches the lowpass-filtered Pan. HECS takes the square root of the LS squared intensity, whose CD would be approximately equal to the square root of the CD of the regressions of squares, and hence greater for the intensity of HECS. Approximately, because the exact CD depends on the distribution of the random variables involved in the regression. With reference to Figure 1b,d, we can conclude that HECS is slightly better than BT-H, because the intensity component lying on a hyper-ellipsoid matches the lowpass-filtered Pan better than the intensity component lying on a hyperplane.

Eventually, the computational cost of all methods employing multivariate regressions (GSA, BT-H, AWLP-H and HECS) is practically identical, because all methods exploit fast algorithms [1,6] in which only the intensity component is calculated and the inverse transformation, also in the nonlinear case of HECS, does not require inversion, thanks to the multiplicative injection gains [26]. Out of the methods compared, SR-D [21] is the slowest, TV [20] intermediate, and all the others are fast, if we do not consider the training time of A-PNN-FT [22].

7. Concluding Remarks

In this paper, we have presented an enhanced version of a popular CS pansharpening method, the HCS fusion technique. The proposed method is insensitive to the format of the data [56], either spectral radiance values or packed digital numbers (DNs), thanks to the use of a multivariate linear regression between the squared interpolated MS bands and the squared lowpass-filtered Pan, in order to find out the MMSE intensity component, which is no longer a linear combination of the interpolated bands. The regression of squares, instead of the Euclidean radius of HCS, makes the color space hyper-ellipsoidal instead of hyperspherical. Furthermore, before fusion is performed, the interpolated MS bands are corrected for the atmospheric path radiance, in order to build a multiplicative injection model with approximately de-hazed components, thus benefiting from the haze correction, crucial for methods exploiting multiplicative detail-injection models [14,15,31,35].

Experiments on true GeoEye-1 and WorldView-3 data show significant advantages over the baseline HCS and its improvements achieved over time. A performance superior to some of the most advanced methods in the literature, including some new-generation methods based on variational optimization, either model based [20,21] or not [22], is achieved. The proposed fusion based on a hyper-ellipsoidal color space (HECS) retains the computational benefits of HCS and the robustness to local misregistration typical of all CS methods.

Author Contributions: Conceptualization and methodology: A.A., L.A., A.G. and S.L.; validation and software: A.A. and A.G.; data curation: A.G.; writing: L.A. and S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Image data supporting the reported results are part of the dataset collection freely distributed by Maxar at <https://resources.maxar.com/product-samples/pansharpening-benchmark-dataset> (accessed on 3 July 2022).

Acknowledgments: The authors wish to give full credit and recognition to the invaluable support of their former coauthor, Bruno Aiazzi, prematurely passed away in December 2021, with whom the present study was originally conceived.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A. *Remote Sensing Image Fusion*; CRC Press: Boca Raton, FL, USA, 2015.
2. Aiazzi, B.; Alparone, L.; Arienzo, A.; Garzelli, A.; Zoppetti, C. Monitoring of changes in vegetation status through integration of time series of hyper-sharpened Sentinel-2 red-edge bands and information-theoretic textural features of Sentinel-1 SAR backscatter. In Proceedings of the Image and Signal Processing for Remote Sensing XXV, Strasbourg, France, 9–11 September 2019; Bruzzone, L., Ed.; SPIE: Bellingham, WA, USA, 2019; Volume 11155, pp. 111550Z–1–111550Z–12.
3. Aiazzi, B.; Alparone, L.; Baronti, S. Information-theoretic heterogeneity measurement for SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 619–624.
4. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. Coherence estimation from multilook incoherent SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2531–2539.
5. D’Elia, C.; Ruscino, S.; Abbate, M.; Aiazzi, B.; Baronti, S.; Alparone, L. SAR image classification through information-theoretic textural features, MRF segmentation, and object-oriented learning vector quantization. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 1116–1126.
6. Vivone, G.; Dalla Mura, M.; Garzelli, A.; Restaino, R.; Scarpa, G.; Ulfarsson, M.; Alparone, L.; Chanussot, J. A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 53–81.
7. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1301–1312.
8. Arienzo, A.; Alparone, L.; Aiazzi, B.; Garzelli, A. Automatic fine alignment of multispectral and panchromatic images. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Waikoloa, HI, USA, 26 September–2 October 2020; pp. 9324689–228–9324689–231.
9. Garzelli, A.; Nencini, F.; Alparone, L.; Baronti, S. Multiresolution fusion of multispectral and panchromatic images through the curvelet transform. In Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Seoul, Korea, 25–29 July 2005; pp. 2838–2841.
10. Aiazzi, B.; Alparone, L.; Baronti, S.; Carlà, R.; Garzelli, A.; Santurri, L. Sensitivity of pansharpening methods to temporal and instrumental changes between multispectral and panchromatic data sets. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 308–319.
11. Aiazzi, B.; Alparone, L.; Garzelli, A.; Santurri, L. Blind correction of local misalignments between multispectral and panchromatic images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1625–1629.
12. Alparone, L.; Garzelli, A.; Vivone, G. Intersensor statistical matching for pansharpening: Theoretical issues and practical solutions. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4682–4695.
13. Vivone, G.; Marano, S.; Chanussot, J. Pansharpening: Context-based generalized Laplacian pyramids by robust regression. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6152–6167.
14. Li, H.; Jing, L. Improvement of a pansharpening method taking into account haze. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 5039–5055.
15. Lolli, S.; Alparone, L.; Garzelli, A.; Vivone, G. Haze correction for contrast-based multispectral pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2255–2259.
16. Pacifici, F.; Longbotham, N.; Emery, W.J. The importance of physical quantities for the analysis of multitemporal and multiangular optical very high spatial resolution images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6241–6256.
17. Garzelli, A.; Nencini, F. Fusion of panchromatic and multispectral images by genetic algorithms. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Denver, CO, USA, 31 July–4 August 2006; pp. 3810–3813.
18. Adesso, P.; Longo, M.; Restaino, R.; Vivone, G. Sequential Bayesian methods for resolution enhancement of TIR image sequences. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**, *8*, 233–243.
19. Fasbender, D.; Radoux, J.; Bogaert, P. Bayesian data fusion for adaptable image pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1847–1857.

20. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O. A new pansharpening algorithm based on total variation. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 318–322.
21. Vicinanza, M.R.; Restaino, R.; Vivone, G.; Dalla Mura, M.; Chanussot, J. A pansharpening method based on the sparse representation of injected details. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 180–184.
22. Masi, G.; Cozzolino, D.; Verdoliva, L.; Scarpa, G. Pansharpening by convolutional neural networks. *Remote Sens.* **2016**, *8*, 1–22.
23. Ma, J.; Yu, W.; Chen, C.; Liang, P.; Guo, X.; Jiang, J. Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Inform. Fusion* **2020**, *62*, 110–120.
24. Ma, J.; Yu, W.; Liang, P.; Li, C.; Jiang, J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Inform. Fusion* **2019**, *48*, 11–26.
25. Padwick, C.; Deskevich, M.; Pacifici, F.; Smallwood, S. WorldView-2 pan-sharpening. In Proceedings of the ASPRS 2010 Annual Conference, San Diego, CA, 27 April 2010; pp. 1–14.
26. Tu, T.M.; Hsu, C.L.; Tu, P.Y.; Lee, C.H. An adjustable pan-sharpening approach for IKONOS/QuickBird/GeoEye-1/WorldView-2. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2012**, *5*, 125–134.
27. Aiazzi, B.; Alparone, L.; Arienzo, A.; Garzelli, A.; Lolli, S. Fast multispectral pansharpening based on a hyper-ellipsoidal color space. In Proceedings of the Image and Signal Processing for Remote Sensing XXV, Strasbourg, France, 9–11 September 2019; Bruzzone, L., Ed.; SPIE: Bellingham, WA, USA, 2019; Volume 11155, pp. 1115507–1–1115507–12.
28. Vivone, G.; Alparone, L.; Chanussot, J.; Dalla Mura, M.; Garzelli, A.; Licciardi, G.A.; Restaino, R.; Wald, L. A critical comparison of pansharpening algorithms. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 191–194.
29. Licciardi, G.; Vivone, G.; Dalla Mura, M.; Restaino, R.; Chanussot, J. Multi-resolution analysis techniques and nonlinear PCA for hybrid pansharpening applications. *Multidim. Syst. Signal Process.* **2016**, *27*, 807–830.
30. Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener. U.S. Patent # 6,011,875, 4 January 2000.
31. Vivone, G.; Alparone, L.; Garzelli, A.; Lolli, S. Fast reproducible pansharpening based on instrument and acquisition modeling: AWLP revisited. *Remote Sens.* **2019**, *11*, 2315.
32. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. Advantages of Laplacian pyramids over “à trous” wavelet transforms for pansharpening of multispectral images. In Proceedings of the Image and Signal Processing for Remote Sensing XVIII; Edinburgh, United Kingdom, 24–27 September 2012; Bruzzone, L., Ed.; SPIE: Bellingham, WA, USA, 2012; Volume 8537, pp. 853704–1–853704–10.
33. Aiazzi, B.; Selva, M.; Arienzo, A.; Baronti, S. Influence of the system MTF on the on-board lossless compression of hyperspectral raw data. *Remote Sens.* **2019**, *11*, 791.
34. Gillespie, A.R.; Kahle, A.B.; Walker, R.E. Color enhancement of highly correlated images-II. Channel ratio and “Chromaticity” Transform techniques. *Remote Sens. Environ.* **1987**, *22*, 343–365.
35. Garzelli, A.; Aiazzi, B.; Alparone, L.; Lolli, S.; Vivone, G. Multispectral pansharpening with radiative transfer-based detail-injection modeling for preserving changes in vegetation cover. *Remote Sens.* **2018**, *10*, 1308.
36. Lolli, S. Is the air too polluted for outdoor activities? Check by using your photovoltaic system as an air quality monitoring device. *Sensors* **2021**, *21*, 6342.
37. Vivone, G.; Arienzo, A.; Bilal, M.; Garzelli, A.; Pappalardo, G.; Lolli, S. A dark target Kalman filter algorithm for aerosol property retrievals in urban environment using multispectral images. *Urban Climate* **2022**, *43*, 101135.
38. Chavez, P.S., Jr. Image-based atmospheric corrections—Revisited and improved. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1025–1036.
39. Chavez, P.S., Jr. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* **1988**, *24*, 459–479.
40. Aiazzi, B.; Alparone, L.; Barducci, A.; Baronti, S.; Pippi, I. Estimating noise and information of multispectral imagery. *Optical Engin.* **2002**, *41*, 656–668.
41. Campbell, J.; Ge, C.; Wang, J.; Welton, E.; Bucholtz, A.; Hyer, E.; Reid, E.; Chew, B.; Liew, S.C.; Salinas, S.; et al. Applying advanced ground-based remote sensing in the Southeast Asian maritime continent to characterize regional proficiencies in smoke transport modeling. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 3–22.
42. Lolli, S.; Di Girolamo, P.; Demoz, B.; Li, X.; Welton, E. Rain evaporation rate estimates from dual-wavelength lidar measurements and intercomparison against a model analytical solution. *J. Atmos. Ocean. Technol.* **2017**, *34*, 829–839.
43. Fu, Q.; Liou, K.N. On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci.* **1992**, *49*, 2139–2156.
44. Lolli, S.; Di Girolamo, P. Principal component analysis approach to evaluate instrument performances in developing a cost-effective reliable instrument network for atmospheric measurements. *J. Atmos. Ocean. Technol.* **2015**, *32*, 1642–1649.
45. Lolli, S.; Sauvage, L.; Loac, S.; Lardier, M. EZ LidarTM: A new compact autonomous eye-safe scanning aerosol Lidar for extinction measurements and PBL height detection. Validation of the performances against other instruments and intercomparison campaigns. *Opt. Pura Apl.* **2011**, *44*, 33–41.

46. Lolli, S.; D'Adderio, L.; Campbell, J.; Sicard, M.; Welton, E.; Binci, A.; Rea, A.; Tokay, A.; Comerón, A.; Barragan, R.; et al. Vertically resolved precipitation intensity retrieved through a synergy between the ground-based NASA MPLNET lidar measurements, surface disdrometer datasets and an analytical model solution. *Remote Sens.* **2018**, *10*, 1102.
47. Wald, L.; Ranchin, T.; Mangolini, M. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 691–699.
48. Aiazzi, B.; Alparone, L.; Baronti, S.; Carlà, R. Assessment of pyramid-based multisensor image data fusion. In Proceedings of the Image and Signal Processing for Remote Sensing IV; Barcelona, Spain, 21–25 September 1998; Serpico, S.B., Ed.; SPIE: Bellingham, WA, USA, 1998, Volume 3500, pp. 237–248.
49. Aiazzi, B.; Alparone, L.; Argenti, F.; Baronti, S. Wavelet and pyramid techniques for multisensor data fusion: A performance comparison varying with scale ratios. In Proceedings of the Image and Signal Processing for Remote Sensing V; Florence, Italy, 20–24 September 1999; Serpico, S.B., Ed.; SPIE: Bellingham, WA, USA, 1999; Volume 3871, pp. 251–262.
50. Du, Q.; Younan, N.H.; King, R.L.; Shah, V.P. On the performance evaluation of pan-sharpening techniques. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 518–522.
51. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200.
52. Khan, M.M.; Alparone, L.; Chanussot, J. Pansharpening quality assessment using the modulation transfer functions of instruments. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3880–3891.
53. Aiazzi, B.; Alparone, L.; Baronti, S.; Carlà, R.; Garzelli, A.; Santurri, L. Full scale assessment of pansharpening methods and data products. In Proceedings of the Image and Signal Processing for Remote Sensing XX; Amsterdam, The Netherlands, 22–25 September 2014; Bruzzone, L., Ed.; SPIE: Bellingham, WA, USA, 2014; Volume 9244, pp. 924402-1–924402-12.
54. Vivone, G.; Restaino, R.; Chanussot, J. A Bayesian procedure for full resolution quality assessment of pansharpened products. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4820–4834.
55. Arienzo, A.; Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A. Reproducibility of spectral and radiometric normalized similarity indices for multiband images. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Yokohama, Japan, 28 July–2 August 2019; pp. 8898662–839–8898662–842.
56. Arienzo, A.; Aiazzi, B.; Alparone, L.; Garzelli, A. Reproducibility of pansharpening methods and quality indexes versus data formats. *Remote Sens.* **2021**, *13*, 4399.
57. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O.; Benediktsson, J.A. Quantitative quality evaluation of pansharpened imagery: Consistency versus synthesis. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1247–1259.
58. Yuhas, R.H.; Goetz, A.F.H.; Boardman, J.W. Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm. In Proceedings of the Summaries of the Third Annual Jpl Airborne Geoscience Workshop; Pasadena, CA, USA, 1–5 June 1992; pp. 147–149.
59. Wald, L. *Data Fusion: Definitions and Architectures—Fusion of images of Different Spatial Resolutions*; Les Presses de l'École des Mines: Paris, France, 2002.
60. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84.
61. Garzelli, A.; Nencini, F. Hypercomplex quality assessment of multi-/hyper-spectral images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 662–665.
62. Arienzo, A.; Vivone, G.; Garzelli, A.; Alparone, L.; Chanussot, J. Full-resolution quality assessment of pansharpening: Theoretical and hands-on approaches. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 2–35.
63. Zhou, J.; Civco, D.L.; Silander, J.A. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *Int. J. Remote Sens.* **1998**, *19*, 743–757.
64. Vivone, G.; Dalla Mura, M.; Garzelli, A.; Pacifici, F. A benchmarking protocol for pansharpening: Dataset, preprocessing, and quality assessment. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 6102–6118.
65. Garzelli, A.; Nencini, F.; Capobianco, L. Optimal MMSE Pan sharpening of very high resolution multispectral images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 228–236.
66. Vivone, G.; Restaino, R.; Chanussot, J. Full scale regression-based injection coefficients for panchromatic sharpening. *IEEE Trans. Image Process.* **2018**, *27*, 3418–3431.
67. Aiazzi, B.; Alparone, L.; Arienzo, A.; Baronti, S.; Garzelli, A.; Santurri, L. Deployment of pansharpening for correction of local misalignments between MS and Pan. In Proceedings of the Image and Signal Processing for Remote Sensing XXIV; Berlin, Germany, 10–13 September 2018; Bruzzone, L., Ed.; SPIE: Bellingham, WA, USA, 2018; Volume 9643, pp. 1078902-1–1078902-12.