Datasets

Four datasets are distributed for the PRISMA contest, namely FR1, FR2, for pansharpening at full spatial resolution (5 m for the panchromatic band P and 30 m for the hyperspectral image HS), and RR1, RR2 for pansharpening at reduced spatial resolution (30 m for $P_{\downarrow}$ and 180 m for $HS_{\downarrow}$).

The images P and HS of the full-resolution datasets FR1 and FR2 have been obtained by extracting a 12km x 12km portion (2400 x 2400 pixels for P and 400 x 400 x N pixels for HS) from the original 30km x 30km PRISMA acquisition, after accurate co-registration.

The 900x900 $P_{\downarrow}$ image and the co-registered 150x150xN $HS_{\downarrow}$ denote the 30-m resolution panchromatic image and the N-band 180-m resolution HS image on a geographical area of 27 km x 27 km.

$P_{\downarrow}$ has been obtained from the original P by using an ideal anti-aliasing lowpass filter, while $HS_{\downarrow}$ has been produced from the original HS by applying spatial filters matching the sensor's MTFs of the VNIR and SWIR bands.

For easy portability, images are in ENVI format, that is a flat-binary raster file in 16-bit unsigned integer data format and Band Sequential (BSQ) interleave type, with an accompanying ASCII header file. Each HS image contains N bands selected from the original VNIR-SWIR PRISMA bands. The ASCII header also contains the values of the central wavelengths of the HS bands.

The following table reports the main characteristics of the four datasets.

Name	**Spatial size (pixels)**	**Number of HS bands N (BSQ interleave type)**	**Geographical extension**
FR1	P: 2400x2400 HS: 400x400	69	12 km x 12 km
FR2	P: 2400x2400 HS: 400x400	63	12 km x 12 km
RR1	$P_{\downarrow}$: 900x900 $HS_{\downarrow}$: 150x150	59	27 km x 27 km
RR2	$P_{\downarrow}$: 900x900 $HS_{\downarrow}$: 150x150	73	27 km x 27 km

Protocols

Reduced resolution assessment measures the similarity of the fused product to an ideal reference, i.e., the original hyperspectral (HS) image. That is possible by degrading the resolutions of both the original HS and the original panchromatic (PAN), and by performing fusion from those degraded data.

Clearly, the choice of the filter is crucial in this validation protocol. In general, the filter is defined for ensuring the *consistency* property of the pansharpening process. Thus, it is straightforward that the resolution reduction of the HS image should be done by exploiting spatial filters matching the HS sensor's modulation transfer functions (MTFs) [1]. In addition, the filter used to degrade the PAN image should be designed to preserve the details that would have been seen if the image were acquired at the reduced resolution. Accordingly, a common choice is the use of an ideal filter [1].

The more similar is the obtained pansharpened image to the original HS image, the higher is the measured quality. Such a similarity degree can be easily computed through score indexes that compare two multi-band images.

- Spectral angle mapper (SAM). Given two spectral vectors, $\boldsymbol{v}$ and $\hat{\boldsymbol{v}}$, both having N components, in which $\boldsymbol{v} = [v_1, v_2, \cdots, v_N]$ is the reference spectral pixel vector and $\hat{\boldsymbol{v}} = [\hat{v}_1, \hat{v}_2, \cdots, \hat{v}_N]$ is the test spectral pixel vector, SAM denotes the absolute value of the spectral angle between the two vectors:
$$SAM(\boldsymbol{v}, \hat{\boldsymbol{v}}) = \cos^{-1}\left(\frac{< \boldsymbol{v},\ \hat{\boldsymbol{v}} >}{\|\boldsymbol{v}\|_2 \cdot \|\hat{\boldsymbol{v}}\|_2}\right).$$
SAM is usually expressed in degrees and is equal to zero if and only 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 the index over the whole scene.
- ERGAS. The index, whose French acronym stands for *relative dimensionless global error in synthesis*, is a normalized dissimilarity index that offers a global indication of the distortion towards the reference of a test multi-band image:
$$ERGAS = 100\frac{d_h}{d_l}\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{RMSE(n)}{\mu(n)}\right)^2},$$
where d_h/d_l is the ratio between pixel sizes of PAN and HS, $\mu(n)$ is the mean (average) of the n-th band of the reference, and N is the number of bands. Low values of ERGAS indicate high similarity between fused and reference HS data.
- $Q2^n$. It is the multi-band extension of the universal image quality index (UIQI). 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 $\boldsymbol{z} = \boldsymbol{z}\,(m,n)$ and $\hat{\boldsymbol{z}} = \hat{\boldsymbol{z}}\,(m,n)$ 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 the product of three terms:
$$Q2^n = \frac{|\sigma_{\boldsymbol{z},\hat{\boldsymbol{z}}}|}{\sigma_{\boldsymbol{z}}\sigma_{\hat{\boldsymbol{z}}}} \cdot \frac{2\sigma_{\boldsymbol{z}}\sigma_{\hat{\boldsymbol{z}}}}{\sigma_{\boldsymbol{z}}^2 + \sigma_{\hat{\boldsymbol{z}}}^2} \cdot \frac{2|\bar{\boldsymbol{z}}||\bar{\hat{\boldsymbol{z}}}|}{|\bar{\boldsymbol{z}}|^2 + |\bar{\hat{\boldsymbol{z}}}|^2},$$
the first of which is the modulus of the HC correlation coefficient (HCCC) between $\boldsymbol{z}$ and $\hat{\boldsymbol{z}}$. The second and the third terms measure contrast changes and mean bias, respectively, on all bands simultaneously. Statistics are calculated on $N \times N$ 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 if and only if $\boldsymbol{z} = \hat{\boldsymbol{z}}$ for all pixels.

Full resolution assessment infers the quality of the pansharpened image at the scale of the panchromatic image without resorting to a single reference image, which is not available. Consequently, the problem of assessing the quality of pansharpened products at full resolution is intrinsically ill-posed. To solve the problem, new distortion measurements have been introduced, such that they do not depend on the unavailable true high-resolution HS data. The Q^* is defined as:

$$Q^* = \left(1 - D_\lambda^k\right)^\alpha (1 - D_S^*)^\beta,$$

which is composed by the product of D_λ^k and D_S^*, quantifying the spectral and the spatial distortions, respectively, exploiting the weights α and β. The higher the Q^* index, the better the quality of the fused product. The maximum theoretical value of this index is one when both D_λ^k and D_S^* are equal to zero.

The spectral distortion index is calculated as follows [1]:

$$D_\lambda^k = 1 - Q\left(\widehat{\boldsymbol{HS}}_\downarrow, \widetilde{\mathbf{HS}}\right),$$

where $\widehat{\boldsymbol{HS}}_\downarrow$ is the MTF-filtered pansharpened HS image considering a resolution ratio equal to R, $\widetilde{\mathbf{HS}}$ is the original HS image interpolated to the PAN scale (R times lower than the HS scale) and Q is the UIQI averaged along the HS spectral bands.

Instead, the spatial consistency, $1 - D_S^*$, proposed in [2], is defined by the multivariate linear regression modeling the relationship between the original high-resolution PAN and the pansharpened HS bands. The figure of merit of the matching between the above-mentioned images is given by the coefficient of determination that is used to measure the spatial consistency [2].

[1] G. Vivone, M. Dalla Mura, A. Garzelli, R. Restaino, G. Scarpa, M. O. Ulfarsson, L. Alparone, and J. Chanussot, “A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods,” vol. 9, no. 1, pp. 53–81, 2021.

[2] L. Alparone, A. Garzelli, and G. Vivone, “Spatial consistency for full-scale assessment of pansharpening,” in Proc. IGARSS, 2018, pp. 5132–5134.