Panchromatic and multispectral image fusion for remote sensing and earth observation: Concepts, taxonomy, literature review, evaluation methodologies and challenges ahead

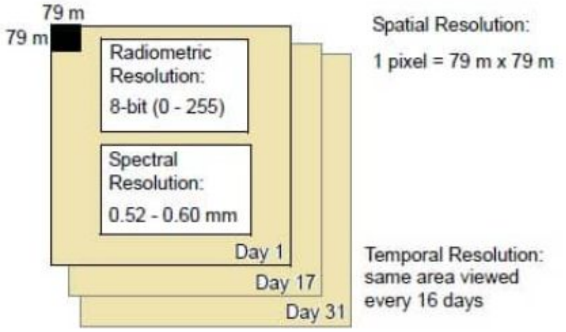
Kai Zhang[A], Feng Zhang[A], Wenbo Wan[A], Hui Yu[B,*], Jiande Sun[A,*], Javier Del Ser[C,D], Eyad Elyan[E], Amir Hussain[F]

[A] School of Information Science and Engineering, Shandong Normal University, China
[B] School of Creative Technologies, University of Portsmouth, UK
[C] TECNALIA, Basque Research and Technology Alliance (BRTA), 48160 Derio, Spain
[D] University of the Basque Country (UPV/EHU), 48013 Bilbao, Spain
[E] School of Computing, Robert Gordon University, UK
[F] School of Computing, Edinburgh Napier University, UK

Group Members:

Pradhyuman Joshi(MT2022074), Prateek Chhimwal(MT2022080), Jaswanth Kumar Chapiri(MT2022051)

Introduction, Motivation, Background to Pansharpening



1. Spatial resolution determines the level of detail in an image.
2. Spectral resolution relates to to the precision in detecting different wavelengths.
3. Radiometric resolution quantifies the sensitivity to intensity variations.
4. Temporal resolution indicates the frequency of data collection over time.

Panchromatic and multispectral image fusion, termed **pan-sharpening**, is to merge the spatial and spectral information of the source images into a fused one, which has a higher spatial and spectral resolution and is more reliable for downstream tasks compared with any of the source images.

Two important attributes of remote sensing images, spatial and spectral resolutions have significant influences on the interpretation accuracy of the observed scene.

Although the spatial and spectral resolutions of remote sensing images are continuously improved, a high spatial and spectral resolution cannot be achieved simultaneously for these images. This is caused by the intrinsic trade-off between spatial and spectral resolutions of imaging sensors.

Problem Statement and Definition

Pan-sharpening is the process of improving the spatial resolution of images by combining a **high-resolution panchromatic** image with one or more **lower-resolution multispectral** images.

The paper discusses different paradigms for pan-sharpening, such as **component substitution**, **multiresolution analysis**, **degradation models**, and **deep neural networks**. It also covers methods for evaluating the quality of the fused images at both reduced-resolution and full-resolution levels.

The survey paper serves as a reference point for newcomers in the field and outlines research directions and challenges in pan-sharpening.

Objective(s) of the paper

The objective of the paper is to provide a comprehensive review of pan-sharpening techniques, with a focus on recent advancements driven by artificial intelligence and deep learning. This includes:

1. **Explaining pan-sharpening**: Clarifying the process of merging panchromatic and multispectral images to improve spatial and spectral resolution.
2. **Reviewing methods**: Evaluating different pan-sharpening approaches categorized into four paradigms.
3. **Image evaluation**: Describing methods to assess the quality of the fused images.
4. **Identifying limitations and challenges**: Discussing issues in the field, including dataset-related challenges.
5. **Development trends**: Summarizing current trends in pan-sharpening, particularly those influenced by AI and DL.
6. **Serving as a reference**: Offering a starting point for newcomers and establishing common research directions for the field.

Methodology

Pan-sharpening methods aim to enhance image quality by merging panchromatic and multispectral images. These methods fall into four categories: component substitution, multiresolution analysis, degradation model-based, and deep neural network-based methods.

1. **CS-based methods**: These methods separate spatial and spectral information in the multispectral (MS) image, up-sample the MS image, create a spatial component, and substitute it with the histogram-matched panchromatic (PAN) image to generate the high-resolution (HR) MS image. Common transformations used include intensity-hue-saturation (IHS), principal component analysis (PCA), and Gram-Schmidt (GS) transformation.
2. **MRA-based methods**: They assume that missing spatial details in the low-resolution (LR) MS image can be inferred from the PAN image. Multiresolution analysis (MRA) tools are used to extract spatial details from the PAN image and inject them into the LR MS image through fusion rules. Variants of MRA are also employed, including support value transformation, support tensor transformation, and morphological filters.
3. **DM-based methods**: In this category, LR MS and PAN images are considered as degradation results of the HR MS image. Pan-sharpening is treated as an image restoration problem, and the fused image is estimated by solving inverse problems derived from spatial and spectral degradation models. Sparsity, gradient priors, and low-rank priors are often used for regularization.
4. **DNN-based methods**: Leveraging the power of deep neural networks (DNNs) and supervised learning, these methods have gained popularity. Various DNNs, including generative adversarial networks and transformers, are used to improve pan-sharpening performance. DNN-based methods can be categorized into three subcategories based on how they combine information from LR MS and PAN images: source image concatenation (SIC), feature concatenation (FC), and feature fusion (FF).

Component Substitution (CS)	IHS (Intensity-Hue-Saturation), PCA (Principal Component Analysis), GS (Gram-Schmidt), BDSD (Band-Dependent Spatial Detail)
Multiresolution Analysis (MRA)	GLP (Generalized Laplacian Pyramid), Wavelet, Nonsubsampled Contourlet and Curvelet, Framelet and Shearlet, MRA-like methods
Degradation Model (DM)	Sparsity Prior, Gradient Prior, Other Priors
Deep Neural Networks (DNN)	Source Image Concatenation (SIC), Feature Concatenation (FC), Feature Fusion (FF), Hybrid methods, Transformer-based methods, Optimization-driven methods

Tabular Representation

Component Substitution-Based Methods

1. IHS (Intensity-Hue-Saturation): These methods use IHS to generate the intensity component of the low-resolution multispectral (LR MS) image, which is then replaced by the panchromatic (PAN) image. They can face limitations when dealing with MS images containing only 3 bands. Adaptive IHS (AIHS) and Improved Adaptive IHS (IAIHS) methods have been developed to calculate weights for each band, reducing spectral distortions in the fused image.
2. PCA (Principal Component Analysis): PCA-based pan-sharpening selects the first principal component (1st PC) of the LR MS image, which is replaced by the PAN image. Adaptive PCA methods aim to find the most similar PC to the PAN image. Variations combine spatial and spectral PCA and apply modulation transfer function (MTF) based filters for improved spatial details.
3. GS (Gram-Schmidt): GS-based methods create orthogonal vectors from the up-sampled LR MS image. The spatial component is replaced by the PAN image, and the fused image is reconstructed. Researchers have attempted various methods to estimate combined weights, but the inflexible orthogonal projection limits room for improvement, resulting in fewer GS-based variants.
4. BDSD (Band-Dependent Spatial Detail): BDSD methods employ content-dependent weights for less spectral distortion. Weight estimation is performed from the corresponding downsampling PAN and MS images. Several variants of BDSD, including C-BDSD, CR-BDSD, and others, further enhance spatial consistency and robustness through different techniques, such as clustering and outlier removal.

In summary, these methods represent various approaches to pan-sharpening, each with unique techniques and variations designed to improve the quality and accuracy of the fused images.

Multiresolution analysis-based methods

In the context of pan-sharpening, CS-based methods have limitations, primarily due to the direct replacement of the spatial component in the low-resolution multispectral (LR MS) image with the panchromatic (PAN) image, resulting in spectral distortions.

To address this issue,Multiresolution Analysis (MRA) is introduced as an alternative approach. In MRA-based methods, only the spatial details from the PAN image are injected into the LR MS image, which helps preserve spectral information. Various tools derived from MRA are employed in pan-sharpening:

1. Generalized Laplacian Pyramid (GLP): Used for high-frequency extraction, GLP methods adjust injection gains at different scales to ensure consistent results.
2. Wavelet: Wavelet-based methods incorporate spectral response and physical properties of observed scenes, improving fusion results. Variants within this framework aim to enhance injection coefficients among bands.
3. Nonsubsampled Contourlet and Curvelet: These advanced MRA tools are employed to capture spatial details and edge representations, resulting in improved pan-sharpening outcomes. Different fusion rules and techniques are applied to enhance spatial information.
4. Framelet and Shearlet: Framelet and shearlet transform methods favor sparse representations and flexible direction feature representation, leading to improved image fusion. Various approaches such as pulse-coupled neural networks and local energy fusion are applied.

MRA-like methods, inspired by MRA, are also developed. These include support value transformation, support tensor transformation, and morphological filters. Efficient filters play a crucial role in these methods.

MRA-based methods excel in preserving spectral information, as only high frequencies are injected into the LR MS image. However, they are sensitive to spatial correspondence and may introduce local dissimilarities or artifacts when the spatial information from the PAN image does not match that of the LR MS image.

Degradation model-based methods

DM-based methods in pan-sharpening aim to efficiently preserve both spatial and spectral information by jointly restoring High-Resolution Multispectral (HR MS) images from Low-Resolution Multispectral (LR MS) and Panchromatic (PAN) images. The spatial and spectral degradation models define the relationship between these images but face an ill-posed problem with multiple potential solutions. Therefore, various priors are introduced to regularize the fusion model.

1. Sparsity Prior: Inspired by compressed sensing, sparsity is a widely used prior in pan-sharpening. It employs the spatial and spectral degradation matrices as measurement matrices and solves the problem using basis pursuit algorithms. Various versions of this framework have been introduced, utilizing HR dictionaries, texture domain priors, and convolution sparse coding to enhance performance.
2. Gradient Prior: Gradient-based priors, like total variation (TV), have been introduced to improve spatial information preservation. Researchers explore different formulations and norms for TV, considering probability distributions of images and investigating the gradient domain relationships between source images and HR MS images. Methods minimize gradient differences to ensure better alignment between PAN and HR MS gradients.
3. Other Priors: Low-rank properties, local, nonlocal, and non-negative priors are also considered to regularize fusion models. These priors help ensure better preservation of spatial information and structural correlations among image components. Methods may combine low-dimensional constraints with other priors to enhance fusion results.

Despite achieving satisfactory HR MS images, DM-based methods have challenges, including high computational complexity due to iterative optimization and a heavy dependence on specific priors. Furthermore, changes in Modulation Transfer Functions (MTFs) of aging imaging sensors can introduce reconstruction errors due to misestimated spatial and spectral degradation matrices.

Deep Neural network-based model

In the past decade, Deep Neural Networks (DNNs) have seen significant success in various fields. They have also made their way into pan-sharpening, where the goal is to fuse Low-Resolution Multispectral (LR MS) and Panchromatic (PAN) images. DNN-based methods can be categorized into three sub-categories based on how they handle the dual-source input:

1. Source Image Concatenation (SIC): This approach directly concatenates LR MS and PAN images and feeds them into DNNs. Various network architectures are utilized, including residual networks, U-Net, multiscale networks, and attention networks. However, SIC may not efficiently exploit complementary information, and abundant data may not be fully utilized.
2. Feature Concatenation (FC): FC deals with information in the feature domain, where feature maps from LR MS and PAN images are combined. This approach may use hand-crafted transformations, such as Gabor filters and Laplacian pyramids, or incorporate sub-networks for feature extraction. While FC is more efficient in preserving information, it may increase computational complexity due to concatenation operations.
3. Feature Fusion (FF): FF aims to merge complementary information from LR MS and PAN images within the feature domain. Various fusion rules are employed, such as addition, subtraction, and multiplication between feature maps at different network levels. FF effectively eliminates redundancy among feature maps, but it can be challenging to devise optimal fusion strategies.

Hybrid methods that combine SIC, FC, and FF are also explored to leverage spatial and spectral information effectively. Additionally, DNNs are rapidly evolving, and recent advancements include the use of transformers and pan-sharpening driven by optimization models. While DNNs have revitalized pan-sharpening, they also present challenges, particularly regarding generalization and model interpretability.

The development of DNN-based methods has significantly advanced the field of pan-sharpening, offering a range of techniques and architectures for improving image fusion.

Sources of Data

The "PanCollection" dataset is a collection of pansharpening datasets from various satellite sensors, including WorldView-2, WorldView-3, QuickBird, and Gaofen-2.

GitHub Link: https://github.com/liangjiandeng/PanCollection

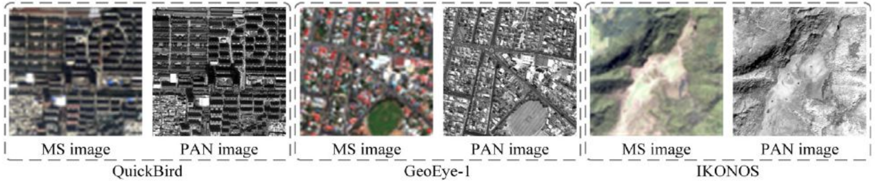
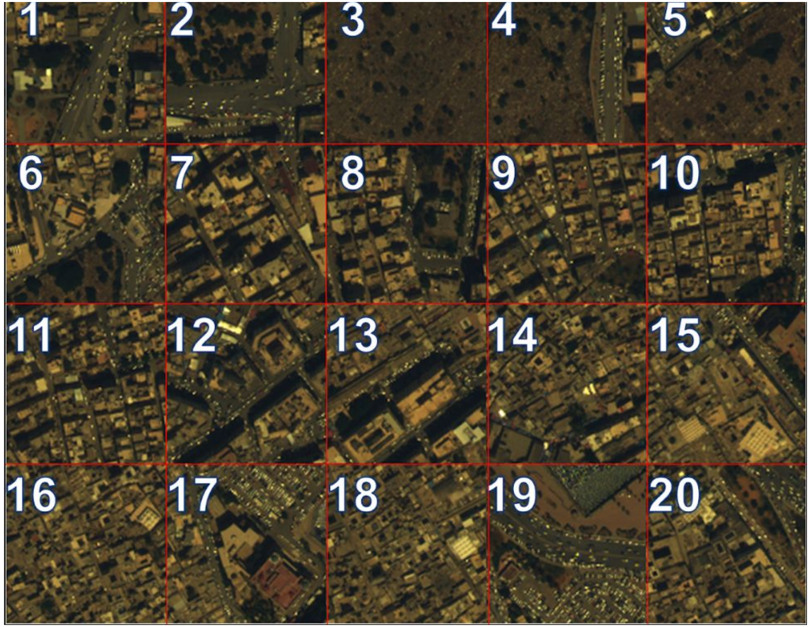


Fig. 1. Multispectral (MS) and panchromatic (PAN) image pairs from different satellites.

Data Overview for WorldView 3 Satellite



Limitations, difficulty and challenges

1. Dataset:
    - As pan-sharpening methods increasingly adopt deep learning techniques, there is a growing need for large and diverse datasets for training these models.
    - However, constructing an all-encompassing dataset for pan-sharpening is challenging due to the diversity of observed scenes, seasonal variations, and differences in imaging sensors.
    - Challenges in dataset construction include accommodating various land cover types (e.g., grassland, woodland, urban areas), accounting for seasonal changes, and addressing the differences between satellite sensors in terms of spectral responses and modulation transfer functions (MTFs).
    - Larger datasets that cover these diverse scenarios are crucial for training DNN-based pan-sharpening methods, improving their generalization capabilities, and achieving better fusion results.

Limitations, difficulty and challenges

2. Quality evaluation:

- Pan-sharpening aims to enhance the spatial resolution of multispectral images. Still, the ultimate goal is to enable meaningful interpretation and downstream tasks like object detection and image segmentation.
- Evaluation of fused images should consider both image-oriented and task-oriented aspects.
- Image-oriented evaluation involves assessing the quality of fused images, with a focus on full-resolution cases. Existing evaluation metrics may not fully capture the visual performance of remote sensing images, and specific no-reference metrics should be designed to address spatial, spectral, and radiometric properties.
- Task-oriented evaluation aims to assess fused images through their impact on downstream tasks. For example, evaluating the effects of different pan-sharpening methods on tasks like change detection.
- Currently, there is a lack of object detection and image segmentation datasets for task-oriented evaluation. The development of such datasets is crucial for making pan-sharpening methods more practical.

Limitations, difficulty and challenges

3. DNNs for pansharpening

- Training Strategies: DNN-based pan-sharpening primarily uses end-to-end training, but challenges arise due to the lack of HR MS training data. Generating training pairs from spatially degraded LR MS and PAN images is common. Exploring training on full-resolution data or combining supervised and unsupervised methods is ongoing.
- New Paradigms: Unsupervised DNNs often lag behind supervised methods. Spatial and spectral degradation models assist training but introduce limitations. Unrolling techniques are promising for integrating spatial and spectral observations. Evolving DNN architectures like transformers, graph neural networks, and zero-reference GANs offer potential solutions for enhancing pan-sharpening quality.

Scope for Future Work

1. ***Reduced-Resolution Metrics:***

   ***Q4 (Quality 4):** Measures how well the colors in the enhanced image match the original high-resolution image.*

   ***SAM (Spectral Angle Mapper):** Calculates the similarity between colors in the enhanced and high-resolution images.*

   ***ERGAS (Erreur Relative Globale Adimensionnelle de Synthèse):** Evaluates the overall error in color synthesis for the enhanced image.*

2. ***Full-Resolution Metrics:***

   ***Dλ (Discrete Lambda):** Quantifies the spectral distortion in the enhanced image compared to the high-resolution image.*

   ***Ds (Spectral Distortion):** Assesses the spectral quality of the enhanced image by comparing it to the high-resolution image.*

   ***QNR (Quality to Noise Ratio):** Evaluates the overall quality of the enhanced image while considering image quality and noise levels.*

Pan-sharpening methods enhance LR MS and PAN image fusion, benefiting interpretation tasks like target detection. These methods fall into four categories: CS-based, MRA-based, DM-based, and DNNs-based. They use various approaches for optimal fusion.

Quality evaluation is conducted using reduced-resolution and full-resolution metrics. In reduced-resolution cases, references like Q4, SAM, and ERGAS are common. Full-resolution metrics like Dλ, Ds, and QNR are used, but more suitable ones need exploration.

Future directions include the need for comprehensive datasets, better image evaluation metrics for full-resolution cases, and the exploration of new DNN paradigms.

Conclusion & Learning Outcome

1. ***Remote sensing*** *is the process of collecting data about Earth's surface or the environment using* ***sensors*** *on satellites or aircraft to study and monitor changes and phenomena* ***from a distance****.*

2. ***Land cover classification*** *is the categorization and mapping of different land cover types, such as* ***forests****,* ***urban areas****, and* ***water bodies****, using satellite or aerial imagery for applications like* ***urban planning*** *and* ***environmental monitoring****.*

3. ***Spectral modeling*** *refers to the analysis and utilization of spectral properties such as* ***reflectance****,* ***absorption****, and* ***spectral bands*** *in remote sensing data for various applications.*

- Pan-sharpening enhances spatial resolution in multispectral imagery for remote sensing applications.
- Fusion of LR MS and PAN images balances spatial and spectral information, aiding tasks like land cover classification and environmental monitoring.
- The field has evolved, incorporating optimization, deep learning, and spectral modeling.
- Offers potential for innovation and addressing challenges in fusion methods, metrics, and dataset diversity.
- We support pan-sharpening projects for their practical significance and research opportunities.