IMPROVING SPATIAL QUALITY OF TERRESTRIAL AND SATELLITE IMAGES BY DEMOSAICKING AND FUSION

by

Fatemeh Fathollahikalanpa

M.Sc. in Remote Sensing, University of Tehran, Iran, 2013

B.Sc. in Geomatics Engineering, University of Tehran, Iran, 2009

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

in the Graduate Academic Unit of Geodesy and Geomatics Engineering

Supervisor: Yun Zhang, PhD, Geodesy and Geomatics Engineering

Examining Board: Shabnam Jabari, PhD, Geodesy & Geomatics Engineering
Rakesh Mishra, PhD, Geodesy & Geomatics Engineering
Julian Meng, PhD, Electrical and Computer Engineering

External Examiner: Jonathan Li, PhD, Geography & Environmental Management,
University of Waterloo

This dissertation is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

May, 2022

ABSTRACT

Improving the spatial quality of a colour image brings valuable benefits to all imaginable applications of the image. One method for such an improvement is to incorporate 'panchromatic' sensors into imaging. Panchromatic sensors provide images with higher spatial quality than colour images because they do not filter any complementary colours of the incoming light. Combining panchromatic and colour sensors has been employed in different fields. In remote sensing (RS), panchromatic and multispectral/hyperspectral images are captured by two separate sensor chips and then fused through pan-sharpening techniques. In terrestrial applications, a single sensor chip is used to accommodate both panchromatic (or white, W) and colour (RGB) pixels using a Colour Filter Array (RGBW CFA). A 'demosaicking' procedure needs to be employed to generate RGB colour images.

Both pan-sharpening and RGBW demosaicking still have unsolved problems despite being used by the imaging industry for a while. In RS, most hyperspectral bands are not pan-sharpened, because they fall beyond the panchromatic spectral range, causing significant spectral distortion. For RGBW demosaicking, limited methods have been published which produce low-quality images, mainly because they demosaick panchromatic and colour images independently. Another issue is that existing approaches cannot handle images corrupted by noise, because they do not involve denoising.

This dissertation aims to overcome the above-mentioned obstacles in improving the spatial quality of the hyperspectral/colour images. For hyperspectral images, this research develops an adaptive band selection strategy to identify the bands across the entire spectrum that can be pan-sharpened without introducing high spectral distortion. For

RGBW demosaicking, this research firstly proposes a collaborative interpolation between panchromatic and colour pixels. It significantly improves the spatial quality by reducing the zipper effects and retaining the spatial details. The research then proposes a deep learning-based approach for RGBW joint demosaicking and denoising, along with a procedure to prepare the required training dataset. Results show a considerable quality improvement over existing methods even for images corrupted by various noise levels. In summary, this research leads to improving the spatial quality of those hyperspectral bands, that were previously left unfused. It also increases the potential of using RGBW cameras in daily applications due to the significant quality boost.

DEDICATION

TO MY BELOVED FAMILY

FOR THEIR SUPPORT, ENCOURAGEMENT, AND INSPIRATION

ACKNOWLEDGEMENTS

I would like to use this opportunity to express my gratitude to everyone who helped make this effort possible. First and foremost, I would like to thank my supervisor, Prof. Yun Zhang, for his trust, advice, support, as well as the additional assistance he gave throughout my PhD program. Second, I would want to express my gratitude to Dr. Shabnam Jabari for her encouragement, suggestions, and assistance in gathering imagery data. During my PhD, David Fraser and other colleagues of the CRC-Lab in Advanced Geomatics Image Processing gave invaluable support. I would like to express my gratitude to them and all the faculty members and staff at the University of New Brunswick's Geodesy and Geomatics Engineering department for all of their help and support. I also would want to thank INO (Canada’s National Optics Institute) for providing technical support and taking industry-standard images for the research conducted in chapters 2 and 3 of this dissertation. I would like to thank Compute Canada (the organization responsible for digital research infrastructure in Canada), and ACENET (the regional partner in Atlantic Canada) for providing Advanced computing resources for the research conducted in chapter 4. Finally, I would want to thank Dr. Mohammad Rezaee, my husband, for his unwavering love, support, patience, and understanding, as well as all of his help and advice as a colleague.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

RGBW	- Red-Green-Blue-White
SNR	- Signal-to-Noise Ratio
RS	- Remote Sensing
CFA	- Colour Filter Array
R	- Red
G	- Green
B	- Blue
W	- White
RGB	- Red-Green-Blue
CFA2.0	- Colour Filter Array 2
CNN	- Convolutional Neural Networks
ARI	- Adaptive Residual Interpolation
PSNR	- Peak Signal to Noise Ratios
HS	- Hyperspectral
MS	- Multispectral
Pan	- Panchromatic
GFM	- General Fusion Model
ALI	- Advanced Land Imager
PCA	- Principal Component Analysis
GS	- Gram Schmid
GSA	- Gram Schmid - Adaptive
SFIM	- Smoothing Filter-based Intensity Modulation

GLP-HPM	- Generalised Laplacian Pyramid with High Pass Modulation
CNMF	- Coupled Non-negative Matrix Factorization
HySure	- Hyperspectral Super-resolution
CS	- Component Substitution
MRA	- Multi-Resolution Analysis
LPF	- Low Pass Filter
VCA	- Vertex Component Analysis
ALI	- Advanced Land Imager
NIR	- Near-Infrared
SAM	- Spectral Angle Mapper
ERGAS	- Erreur Relative Globale Adimensionnelle de Synth`ese
QNR	- Quality with No Reference
RMSE	- Root Mean Square Error
MTF	- Modulation Transfer Function
UIQI	- Universal Image Quality Index
CI	- Collaborative Interpolation
QE	- Quantum Efficiency
HRP	- High Resolution Panchromatic
LRC	- Low Resolution Colour
MLRI	- Minimized-Laplacian Residual Interpolation
Tik	- Tikhonov
TV	- Total Variation
CPSNR	- Colour Peak Signal-To-Noise Ratio

LRP	- Low Resolution Panchromatic
RI	- Residual Interpolation
BC	- Bicubic
GF	- Guided Filtering
OD	- Optical Density
CD	- Channel Discrepancy
UAV	- Unmanned Aerial Vehicle
P	- Panchromatic
ResNet	- Residual Networks
RGBW-DJDD	- RGBW Deep Joint Demosaicking and Denoising, or briefly RGBW-DJ
VAR	- Variational Approach
ETS	- (Condat, 2009)Effective Three-Stage method
BN	- Batch Normalization
ELU	- Exponential Linear Unit
WED	- Waterloo Exploration Dataset
DL	- Deep Learning
DMCNN_VD	- Demosaicking CNN_ Very Deep

1 Introduction

This PhD dissertation leads to strategies to improve the image spatial quality in two different fields; the hyperspectral images in remote sensing, and the colour images obtained from RGBW demosaicking in terrestrial applications.

1.1 Background

Improving the spatial quality of the colour images greatly expands their use and applicability, regardless of the platform on which they are captured. In this dissertation, spatial quality refers to the quality of the fine details in an image in terms of how observable, distinguishable, and noise free those details are. Thus, improving the spatial quality results in better objects discrimination or recognition by human eyes or by a machine trained to do so. Such improvement significantly increases the accuracy of classification and object detection which are highly common image applications.

Improving spatial quality is ongoing in both hardware and software fields. In the hardware field, it is followed by developing more advanced sensors that provide images with a higher spatial resolution (i.e., the size of the smallest object that a sensor resolves (Liang & Wang, 2020)). This is done by reducing the pixel size of the sensors or, in other words, by increasing the pixel numbers in a given area. However, the manufacturing constraints do not allow us to shrink the pixels as much as we want. Besides, by decreasing the pixel size, the amount of energy received by the pixel, and hence, the signal-to-noise ratio (SNR) decreases (Farrell et al., 2006). This especially challenges low light shooting when the energy source is already reduced. Thus, due to the manufacturing limitations and the rising manufacturing cost of such sensors, as well as the trade-off between the pixel size and the SNR, spatial quality improvement is also pursued in the software field. In the software field, after the images are captured, image processing techniques are employed as a post-processing step by manipulating digital values based on the mathematical models. Such improvements are far less expensive compared to hardware improvements. In addition, regardless of how advanced future sensors become, image processing techniques can be used to provide even higher quality.

Among image processing methods, there are two distinct branches of studies to improve spatial quality; One focuses on injecting the spatial details from other higher quality images into lower quality ones (relevant literature will follow). The other approach, called "super resolution" focuses on improving the spatial resolution without using another image and obtains a high resolution output from one of its low resolution versions (Y. Chen et al., 2021; L. Wang et al., 2021; Yang et al., 2019). This dissertation, however, focuses on the former, i.e., combining a lower quality image with a higher quality one and attempts to overcome some of the remaining challenges.

In the imaging industry, panchromatic sensors are recognized to give superior spatial quality than colour sensors. Panchromatic sensors do not produce colour images, since they receive energy from a wider range of the electromagnetic spectrum than with colour sensors (the whole visible and a portion of the near-infrared range). With a larger amount of energy received per pixel, i.e., with a higher SNR, it is possible to reduce the pixel size, and hence to increase the spatial resolution. As a result, the panchromatic images provide higher spatial quality, which is later integrated into the lower resolution colour images.

Looking into imaging systems, combining a panchromatic and a colour sensor has been employed on a variety of platforms, from satellite to terrestrial, and involves mainly two image processing techniques 'pan-sharpening' and 'demosaicking'. In Remote Sensing (RS), most of the modern satellites carry separate sensor chips that simultaneously capture panchromatic and multispectral images. Pan-sharpening is a term used mostly in RS and refers to a specific branch of 'image fusion' techniques. It involves fusing a multispectral image with a higher resolution panchromatic image. This generates a spatially improved multispectral image, whose spatial information is provided by the panchromatic image. Terrestrial imaging systems, on the other hand, mostly use single-chip cameras equipped with a Colour Filter Array (CFA) on top of the sensor, which has red (R), green (G), and blue (B) filter elements. To improve the sensor SNR, panchromatic filter elements (also known as white, W, in the imaging industry) are inserted into a CFA that generates a newer version called RGBW CFAs. A single-chip sensor captures a single-channel raw image that must be processed via demosaicking (or debayering) methods to reconstruct a full resolution three-channel colour image. The demosaicking method should be designed in such a way that the panchromatic spatial information could be effectively transferred into the colour image.

Even though combining panchromatic and colour sensors has been done in RS for more than two decades and in RGBW CFAs for more than a decade, there are still unsolved issues in both fields. The significant studies that yielded to identifying such issues are briefly reviewed in the following section.

1.2 Literature Review

In this section, the literature review is categorized based on the two areas of interest: 1) Hyperspectral pan-sharpening and 2) RGBW demosaicking.

1.2.1 Hyperspectral Pan-sharpening

Compared to multispectral pan-sharpening methods reviewed by Meng et al., (2019), there are few methods dedicated to the spatial enhancement of hyperspectral images (Cetin & Musaoglu, 2009; Z. Chen et al., 2014; Moeller et al., 2009; Simoes et al., 2015; Vivone, Restaino, Licciardi, et al., 2014; Wei, 2015; Yokoya et al., 2012). According to a general rule in pan-sharpening, the spectral bands to be fused should cover the same spectral range as the panchromatic band. Otherwise, the resultant image would suffer from huge spectral distortion (Cetin & Musaoglu, 2009). Thus, the main reason that hyperspectral pan-sharpening remains a challenge is the spectral discrepancy between the hyperspectral and panchromatic bands (Loncan et al., 2015). A common practice is to fuse only the bands that fall within the panchromatic spectral range to avoid spectral distortion in the fused images (Cetin & Musaoglu, 2009; Simoes et al., 2015; Vivone, Restaino, Licciardi, et al., 2014). However, with such an approach, only 10% of the hyperspectral bands can be involved in the pan-sharpening process. In other words, most of the hyperspectral bands remain unfused, and no improvement in their spatial quality can be achieved.

Other studies that use airborne hyperspectral sensors have synthesized the panchromatic image by spectrally degrading the hyperspectral bands (Loncan et al., 2015; Wei et al., 2016; Yokoya et al., 2012), for example, by averaging all the hyperspectral bands (Wei et al., 2016). This implies that the common spectral range between panchromatic and hyperspectral bands is synthesized and set as desired. Thus, this cannot the applied when using real satellite images for practical applications.

1.2.1.1 Research Premises

From the literature review above, it can be inferred that the only realistic band selection scenario for hyperspectral pan-sharpening is to select the bands within the panchromatic spectral range. This way, most of the bands that have a wavelength beyond the panchromatic spectral range cannot be pan-sharpened, i.e., they cannot be spatially improved. Thus, the feasibility of fusing bands outside of the panchromatic spectral range has not been investigated and forms the fundamentals of the second chapter of this dissertation.

1.2.2 RGBW Demosaicking

Incorporating panchromatic pixels into RGB CFAs was first patented in 2007 (Hamilton & Compton, 2007). Then, Kodak developed a high-performance sensor equipped with a specific RGBW CFA, called Truesense sparse CFA (AND9180/D-TRUESENSE Sparse Color Filter Pattern, 2014), commonly referred to as CFA2.0 in the literature. It generates a raw image where half of the pixels are panchromatic, and the other half belongs to R, G, and B pixels.

Studying the literature, it is understood that some methods involve two main procedures, i.e. demosaicking and fusion, such as (Chung et al., 2020; Hamilton & Compton, 2007; Kwan et al., 2017; Kwan & Chou, 2019), while others involve only demosaicking, like (Condat, 2009, 2012; Kang & Jung, 2020; Rafinazari, 2017; Singh &

Singh, 2012; J. Wang et al., 2011; C. Zhang et al., 2016). Kodak employs a demosaicking-and-fusion approach, in which colour and panchromatic images are interpolated and then integrated (Hamilton & Compton, 2007). The panchromatic image is interpolated using only panchromatic pixels without involving colour pixels. For the colour image, the panchromatic pixels are used directly for G interpolation and the resulting G is subsequently used for R and B interpolations. However, the outcomes of such an approach suffer from visible colour leakages and zipper artifacts especially in the areas related to high frequency information. In other words, the spatial quality is distorted.

The demosaicking-and-fusion process was used in two other studies, although with different demosaicking and fusion techniques using simulated RGBW images. In (Kwan et al., 2017), the panchromatic and colour images were demosaicked independently during the demosaicking step. For the fusion step, they recommended using the pan-sharpening methods utilized in RS. They came to the conclusion that their preliminary results did not indicate a substantial improvement over Kodak's approach. In the other study (Chung et al., 2020), the panchromatic image was first interpolated using a cross shape-based colour difference approach. For the fusion, they exactly followed the approach in (Hamilton & Compton, 2007). However, the obtained colour image was further enhanced through two more steps; an iterative error compensation-based approach and a refinement process based on the RGBW image.

Other methods which include only demosaicking step, mainly focus on estimating the luminance and chrominance components either in the frequency domain (Rafinazari, 2017; J. Wang et al., 2011) or in the spatial domain through the so-called universal methods (Condat, 2009, 2012; Kang & Jung, 2020; Singh & Singh, 2012; C. Zhang et al., 2016). The evaluations in all the mentioned studies have been performed using simulated RGBW images.

1.2.2.1 Research Premises

Indicated above, a very limited study has been devoted to RGBW demosaicking, and thus, the following study gaps can be found and assessed:

First, the results of the existing methods suffer severely from degraded spatial quality which is reflected as jagged edges (or zipper effects), especially in areas where high-frequency information exists. This means that the existing methods are not very successful in injecting the spatial features of panchromatic pixels into the colour pixels. More precisely, in the demosaicking-and-fusion techniques, the panchromatic and colour pixels were not effectively involved in each other's demosaicking. In fact, the strong spatial correlation between the colour and panchromatic pixels is neglected in these methods, while it could be utilized collaboratively to guide the interpolation of each of them. In addition, most of the existing methods have not been tested using real RGBW images captured by an actual RGBW sensor. In other words, the capability of these methods for real-world application is questionable. Therefore, the possibility of employing a collaborative interpolation between panchromatic and colour pixels, in addition to its applicability for real RGBW images serve as the premises for the research conducted in the third chapter of this dissertation.

Second, the existing studies neglect denoising as one of the crucial steps in the image formation pipeline. None of the mentioned studies conducted a thorough investigation on the effect of different noise levels (i.e., usually only one noise level is investigated). This motivated us to look at incorporating denoising into demosaikcing in a separate chapter (chapter 4 of this dissertation).

1.2.3 RGBW Joint Demosaicking and Denoising

To integrate denoising into RGBW demosaicking, it is reasonable to review current research on RGB CFAs, for which demosaicking and denoising are widely explored. According to two recent studies (Jin et al., 2020; Kokkinos & Lefkimmiatis, 2019), a joint demosaicking and denoising technique works better than any other approach that breaks these two processes into sequential steps. Moreover, by investigating the joint demosaicking and denoising methods for RGB CFAs, it could be seen that they are mostly based on Convolutional Neural Networks (CNN) derived from deep learning (Din et al., 2020; N. Dong et al., 2021; W. Dong et al., 2018; Gharbi et al., 2016; Kokkinos & Lefkimmiatis, 2019; Pan et al., 2020; Syu et al., 2018; H. Tan et al., 2022; Wu et al., 2021). The achievements of the literature review on RGB CFAs motivated us to bring such concepts into RGBW CFAs by aiming at a joint demosaicking and denoising approach via designing a deep CNN model.

However, employing deep CNNs for RGBW demosaicking encounters another main issue, i.e., the shortage of ground-truth panchromatic images required for training the network. The non-learning-based methods use benchmark datasets (mostly Kodak and IMAX) and simulate the panchromatic pixels as the average of the R, G, and B pixels (Condat, 2009, 2012; Kwan et al., 2017; Rafinazari, 2017; Singh & Singh, 2012; J. Wang et al., 2011; C. Zhang et al., 2016). However, such a simulation would create nonrealistic training datasets when used to train a network. The network would learn during the training procedure that the panchromatic pixel values are the average of the colour pixel values. When subjected to a real RGBW image, where the relationship between panchromatic and colour pixels is no longer a simple average, such a network would not function effectively. Therefore, generating the panchromatic images that best simulate the imaging reality is another issue that has not been addressed in the literature.

1.2.3.1 Research Premises

Indicated above, a joint demosaicking and denoising approach has not been studied for RGBW CFAs. Thus, employing a joint demosaicking and denoising approach through a deep CNN for RGBW images, besides its applicability to real images, is to be investigated in the fourth chapter. Furthermore, simulating more realistic panchromatic images to serve as the training dataset is an issue that must be researched and explored.

1.3 Research Questions

Through the literature review, it was learnt that there are still problems in terms of enhancing the spatial quality in different imaging fields, mainly in hyperspectral pan-sharpening and RGBW demosaicking.

For hyperspectral pan-sharpening, the main question is how to incorporate the hyperspectral bands that are beyond the panchromatic spectral range without producing much spectral distortion in the final fused image. Would statistical measurements on panchromatic and hyperspectral bands help to find those bands which spectrally correlate with the panchromatic image? The next question is to see if incorporating those bands will affect the spectral quality of the bands that fall within the panchromatic spectral range. Another question to answer would be which pan-sharpening method is more flexible in accepting the spectral bands which do not match the panchromatic spectral range.

For RGBW demosaicking, the first main question is to see if a collaborative interpolation between colour and panchromatic pixels better improves the spatial quality in a demosaicking-and-fusion approach. For this research, collaborative interpolation will be performed by considering each of the panchromatic/colour images as an auxiliary observation to guide demosaicking the other one. Finally, it will assess wether the proposed approach is practically applicable to real RGBW images.

For RGBW demosaicking, another main question to answer is whether denoising can be involved in RGBW demosaicking. It will be investigated how to adapt a joint demosaicking and denoising for RGBW images through a deep CNN. Another question to be answered is how to prepare an appropriate training dataset that would contain more realistic panchromatic pixels using existing RGB datasets. Finally, it will assess wether the proposed approach is practically applicable to real RGBW images.

1.4 Research Objectives

The main objective of this research is to overcome the weaknesses for improving the spatial quality of the hyperspectral images in RS and the colour images from RGBW CFAs in terrestrial applications. More specifically:

For hyperspectral images, the objective is to study the feasibility of improving the spatial quality of those hyperspectral bands that are beyond the panchromatic spectral range through different pan-sharpening techniques. To do so, it establishes a systematic way to select the appropriate bands outside of the panchromatic spectral range to feed into pan-sharpening methods. This is accomplished by employing the statistical dependencies between panchromatic and hyperspectral images. Such selection should not result in high spectral distortion in the final images. The spectral distortion can be interpreted by comparing spectral quality metrics besides the spectral signatures obtained from fused images and the original hyperspectral images. The second objective is to compare various pan-sharpening methods and find which method leads to the least spectral distortion.

For RGBW images, the first main objective is to provide a demosaicking-and-fusion procedure that produces higher quality demosaicked images compared to the existing methods, as well as the results of the commercial camera equipped with Truesense CFA. Accordingly, the colour and panchromatic images would be demosaicked through a collaborative interpolation, followed by fusing colour and panchromatic images.

Through collaborative interpolation, the colour pixels are incorporated in demosaicking the panchromatic image, and the panchromatic pixels are involved in demosaicking the colour image. To do so, it is proposed to use guided filter technique (He et al., 2013) which is previously implemented in an RGB demosaicking algorithm, Adaptive Residual Interpolation (ARI; Monno et al., 2017). However, modifications to the ARI method are required to make it applicable to the RGBW CFAs. The fusion step is then fulfilled through a pan-sharpening method, UNB-pansharp (Y. Zhang & Mishra, 2014).

The proposed procedure will be evaluated using two scenarios: 1) simulating RGBW images using the benchmarks (as indicated in the literature), and 2) using real RGBW images captured by a Truesense camera (BOBCAT IGV-B2320), which provides its own demosaicked images besides raw images. The latter scenario is missing in the experiments found in the literature, but it is more favourable since simulations cannot fully reflect real-world imaging conditions. The performance of different methods are compared by calculating commonly used metrics, peak signal to noise ratios (PSNR) and CIE-LAB error (Chung et al., 2020).

For RGBW images, the second main objective is to integrate denoising procedure into RGBW demosaicking using CNN models. In particular, this study introduces a network specifically designed for RGBW CFAs inspired by the ideas presented by (Syu et al., 2018) and (Fletcher, 2018). The lateral objective is to simulate the panchromatic images using available datasets that contain RGB channels. To do so, a linear transformation is established to convert RGB images into panchromatic ones. This is inspired by a similar idea in the pan-sharpening application, where the panchromatic image is spectrally related to the colour channels through a linear combination (Fathollahi & Zhang, 2020; Meng et al., 2019). The panchromatic images simulated this way serve as the fourth channel of the available datasets. These 4-channel images are then used as the training and testing datasets to train and test the joint demosaicking and denoising network. The evaluation is accomplished by calculating PSNR and CIE-LAB error.

1.5 Dissertation Structure

This dissertation includes five chapters. The current chapter 1 serves as an introduction to the dissertation, providing a background, literature review, research questions and objectives. Chapters 2 concentrates on the research regarding hyperspectral images and investigates an adaptive band selection for pan-sharpening. Chapter 3 relates to demosaicking the RGBW images using a collaborative interpolation between panchromatic and colour pixels. Chapter 4 presents a deep joint demosaicking and denoising for RGBW images. It is noteworthy that the contents of chapters 2 to 4 present three journal papers, published or under review, as follows:

- Chapter 2: Fathollahi, F., & Zhang, Y. (2020). Adaptive band selection for pan-sharpening of hyperspectral images. International Journal of Remote Sensing, 41(10), 3924–3947.
- Chapter 3: Fathollahi, F., Jabari, S., Zhang, Y., & Rezaee, M. (2021). RGBW Demosaicking Using a Collaborative Interpolation Between Panchromatic and Colour Pixels. Signal Processing: Image Communication. Under review
- Chapter 4: Fathollahi, F., Rezaee, M., & Zhang, Y. (2021). Deep Joint Demosaicking and Denoising for RGBW Colour Filter Arrays. IEEE Transactions on Computational imaging. Under review

Finally, chapter 5 concludes the dissertation which contains the summary of the work, contributions, and recommendations for future work.

2 Adaptive Band Selection for Pan-sharpening of Hyperspectral Images[1]

2.1 Introduction

Hyperspectral (HS) sensors provide a huge number of spectral bands in order to precisely reconstruct the spectra of different objects on Earth, which can be used for their precise determination and discrimination. Technically, providing such narrow spectral bands requires large ground sampling, i.e. lower spatial resolution due to SNR requirements. In contrast, the panchromatic (Pan) band possesses high-spatial resolution but suffers from low-spectral resolution. However, in many remote sensing applications, such as classification, datasets that maintain both high spectral and spatial resolutions increase the quality/accuracy of the results. To obtain such datasets, one practical solution is to fuse a high-spatial resolution Pan band with a low-spatial resolution HS image using pan-sharpening techniques.

In the past two decades, significant research has been devoted to multispectral (MS) pan-sharpening and various fusion methods have been proposed (Xu et al., 2014). These methods have been classified most commonly into two groups, known as component substitution and multi-resolution analysis, (Choi et al., 2011; Picone et al., 2017). Palubinskas and Reinartz (2011) showed that all these methods in different categories can be formulated using a general fusion model, GFM (the details of GFM are discussed in

---

[1] This chapter is a published paper with the same title in the International Journal of Remote Sensing, 41(10), 3924–3947. However, there have been modifications as suggested by the examiners.

Section 2.2.1). Compared to MS pan-sharpening, there are very few studies, which have been dedicated to the spatial enhancement of the HS images, by fusing HS images with either higher resolution MS bands or a Pan band through pan-sharpening procedures (Cetin & Musaoglu, 2009; Z. Chen et al., 2014; Moeller et al., 2009; Simoes et al., 2015; Vivone, Restaino, Licciardi, et al., 2014; Wei et al., 2016; Yokoya et al., 2012). Among these studies, few of them have studied adopting common MS pan-sharpening methods for HS pan-sharpening (Cetin & Musaoglu, 2009; Z. Chen et al., 2014; Loncan et al., 2015; Vivone, Restaino, Licciardi, et al., 2014), while others have developed specific algorithms for fusing HS with MS/Pan images, mainly based on Matrix Factorization approach (Yokoya et al., 2012) or Bayesian methods (Simoes et al., 2015; Wei et al., 2016). A review of these methods can be found in (Loncan et al., 2015).

One of the main challenges in fusing HS images is the spectral discrepancies between HS bands and Pan/MS bands (Loncan et al., 2015). The HS bands cover usually from visible to shortwave infrared range of the electromagnetic spectrum, while the spectral range of a Pan band is usually limited to only the visible range. To prevent fused images being spectrally distorted, a well-known practice is to select the spectral bands that cover the same spectral range as that of the Pan band (Cetin & Musaoglu, 2009). For MS pan-sharpening, such a practice does not cause much of a problem since the spectral range of the Pan band covers all/ most of the MS bands; thus, all the MS bands can be fused. In contrast, when the common spectral range between HS and Pan images is very narrow, most of the HS bands cannot be pan-sharpened without spectral distortion. As an example, in EO-1 datasets, only 10% (20 bands) of the HS bands fall within the Pan spectral range, thus, the final fused image would have only 10% of the HS bands. This would decrease the efficiency of the fused HS image in lateral applications like discriminating the targets with similar spectrums for classification purposes.

Considering the band selection scenarios in literature, studies using real satellite images have selected only the HS bands which overlap the spectral range of the Pan band (Cetin & Musaoglu, 2009; Simoes et al., 2015; Vivone, Restaino, Licciardi, et al., 2014). With such an approach, all the bands covering the areas beyond the Pan range remain unfused. Other studies have synthesized Pan or MS images by spectrally degrading the HS images acquired by airborne sensors (Loncan et al., 2015; Wei et al., 2016; Yokoya et al., 2012). This simulation is usually done by multiplying some selected bands of a reference HS image with a suitably chosen spectral response function (Loncan et al., 2015; Yokoya et al., 2012), or by averaging all the bands of an HS image (Wei et al., 2016). However, when simulating Pan/MS images, what bands to select for simulation is under control. In other words, the common spectral range of Pan/MS bands with HS bands can be arbitrarily determined, e.g. if all bands are used to simulate a Pan band, it synthetically implies that the spectral range of Pan image covers the whole spectral range of the HS image, and thus, all HS bands can be fused. However, as mentioned before, this is not the case when using real satellite images for practical applications, i.e. the MS/Pan spectral ranges do not cover the whole HS spectral range and thus, we cannot fuse all HS bands.

In summary, it can be seen that to fuse HS images with either real satellite Pan/MS images or simulated Pan/MS images, only the HS bands which share common spectral ranges with Pan/MS bands are selected and used for fusion. Recently, Picone et al. (2017) have studied band selection for fusing MS and HS images acquired by EO-1 satellite.

However, they also have focused on the HS bands overlapping the common spectral ranges with MS bands (B11-B57) and have not considered the bands beyond the MS range. This chapter, however, focuses on fusing the HS images with the Pan image and investigates the feasibility of incorporating those HS bands that are beyond the Pan spectral range, in the pan-sharpening procedure, while maintaining the least spectral distortion in the resultant images. To select the appropriate bands and at the same time to keep the spectral distortion as low as possible, this research suggests using those bands which have higher statistical dependencies with the Pan image.

For this experiment, the HS bands of the Hyperion sensor from EO-1 satellite are fused with the Pan image of the Advanced Land Imager (ALI) sensor from the same satellite. The adaptive band selection method is mainly proposed to adapt the classical MS pan-sharpening methods for the HS pan-sharpening since it is mainly designed based on the GFM. However, we have tested the proposed band selection approach for two other pan-sharpening techniques, specifically designed for HS data. The MS pan-sharpening methods tested include Principal Component Analysis (PCA; Chavez & Kwarteng, 1989), Gram Schmid (GS; Laben & Brower, 2000) and its adaptive version (GSA; Bruno Aiazzi, Baronti, & Selva, 2007), UNB-Pansharp (Zhang, 2008), Smoothing Filter-based Intensity Modulation (SFIM; Liu, 2000), Generalised Laplacian Pyramid with High Pass Modulation (GLP-HPM; Vivone, Restaino, Mura, Licciardi, & Chanussot, 2014). we also have extended our experiment to two specifically designed HS pan-sharpening methods, Coupled Non-negative Matrix Factorization (CNMF; Yokoya et al., 2012) and Hyperspectral Super-resolution (named as HySure; Simoes et al., 2015).

In addition, the size of the datasets used for this experiment (around 3000×1500 pixels for the high resolution images) is much larger than the size of the images used by other mentioned papers (typically, 500×500 pixels at most for the high resolution images), which makes the achievements of this study more meaningful and useable for real-world applications. Results clearly show that with the proposed adaptive band selection, it is possible to increase the number of HS bands that can be pan-sharpened without producing high spectral distortion in the final fused images, even if the added bands are beyond the Pan spectral range.

2.2 Methodology

In this section, first we briefly review the classical pan-sharpening methods adapted for HS image, followed by two other methods developed specifically for HS pan-sharpening. Here, we do not discuss all the procedure details of the different methods (for details the reader is referred to Loncan et al. (2015)), but rather focus on the mathematical models in different methods to provide insight on the spectral relationship between Pan and HS bands. Later, we explain the approach of selecting appropriate bands for pan-sharpening.

The main notation used in this chapter is as follow. We denote the observed low resolution HS image by the matrix $\mathbf{X}^{\mathrm{l}} = \left[\boldsymbol{x}_1^{\mathrm{l}}\ \boldsymbol{x}_2^{\mathrm{l}} \ldots \boldsymbol{x}_{n_h}^{\mathrm{l}}\right] \in \mathbb{R}^{m_h \times n_h}$, also referred to as original HS, where each column, $\boldsymbol{x}_i^{\mathrm{l}}$**,** consists of the spectrum related to the i-th pixel, with m_h and n_h referring to the number of spectral bands and pixels, respectively. We also denote the high resolution Pan band by a vector, $\boldsymbol{Y}^{\mathrm{h}} \in \mathbb{R}^{1 \times n_p}$, where n_p refers to the total number of pixels in the Pan image. Accordingly, $\mathbf{X}^{\mathrm{h}} = \left[\boldsymbol{x}_1^{\mathrm{h}}\ \boldsymbol{x}_2^{\mathrm{h}} \ldots \boldsymbol{x}_{n_p}^{\mathrm{h}}\right] \in \mathbb{R}^{m_h \times n_p}$ refers to the high resolution HS image to be estimated through the fusion procedure (or simply to the fused image), and $\widetilde{\mathbf{X}}^{\mathrm{l}} \in \mathbb{R}^{m_h \times n_p}$ represents the upsampled version of the HS image, to the size of the Pan image.

2.2.1 Pan-sharpening Methods Adapted for HS Images

The wide diversity of existing pan-sharpening methods developed for MS images have encouraged the researchers to adapt and apply them for HS pan-sharpening (Loncan et al., 2015; Vivone, Restaino, Licciardi, et al., 2014). These methods can be differently categorized based on different aspects of the view. One popular categorization, which considers the fusion procedure from an image processing aspect divides the pan-sharpening methods into Component Substitution (CS) and Multi-Resolution Analysis (MRA) groups (Choi et al., 2011; Picone et al., 2017). Another aspect of view is to consider the mathematical model of the fusion procedure. With such a perspective, the pan-sharpening methods are grouped into additive and multiplicative sharpening models (Vivone et al., 2015). However, it has been shown that the multiplicative model is, in fact, a particular case of the additive model where a logarithmic transformation of data is considered (Palubinskas & Reinartz, 2011) or by the choice of proper injection gains (Vivone et al., 2015). It also has been proved that by considering the mathematical point of view, the GFM is achievable for all the classical pan-sharpening methods, no matter to what group they belong (Palubinskas & Reinartz, 2011).

In the same manner, the GFM can be used for HS pan-sharpening, i.e., to obtain the i-th spectral band of the high resolution HS image, we can write (Loncan et al., 2015; Picone et al., 2017)

$$\mathbf{X}^{\mathrm{h}} = \widetilde{\mathbf{X}}^{\mathrm{l}} + \mathbf{g} \odot \left(\boldsymbol{Y}^{\mathrm{h}} - \boldsymbol{Y}^{\mathrm{l}}\right). \tag{2.1}$$

Where $\mathbf{g} = [\boldsymbol{g}_1\ \boldsymbol{g}_2\ \dots\ \boldsymbol{g}_{n_p}]$ is the injection gain matrix, and $\boldsymbol{g}_i \in \mathbb{R}^{m_h \times 1}$ is the injection gain vector containing the gains of different spectral bands for the i-th pixel, and $\odot$ denotes the component-wise multiplication, and $\boldsymbol{Y}^{\mathrm{l}} \in \mathbb{R}^{1 \times n_p}$ is a spatially degraded version of the Pan image with the same number of pixels as $\boldsymbol{Y}^{\mathrm{h}}$. In fact, different pan-sharpening methods vary based on how the injection gains and $\boldsymbol{Y}^{\mathrm{l}}$ are calculated.

2.2.1.1 CS methods

The main assumption in these methods is that when the low resolution spectral bands are projected into a proper orthogonal space (e.g. through PCA or GS transformations), most of the spatial information of the bands are gathered in the first component in the new space (e.g. PC_1 or GS_1), which later, can be replaced by the Pan image to inject the high resolution spatial information. Therefore, $\boldsymbol{Y}^{\mathrm{l}}$ in these methods is obtained through a linear transformation of the spectral bands, and the injection gains are constant for all the pixels in the i-th band; i.e. $\mathbf{g} \in \mathbb{R}^{m_h \times 1}$.

PCA: in this method, $\boldsymbol{Y}^{\mathrm{l}}$ is defined to be the first principal component obtained from applying the principal component transformation on the spectral bands, and the injection gain for the i-th band is the i-th eigenvalue in the first eigenvector (Tu et al., 2001). In other words, if $\boldsymbol{\phi}_1 \in \mathbb{R}^{1 \times m_h}$ denotes the first eigenvector of the covariance matrix of the HS bands, and the operator $(.)^{\mathrm{T}}$ represents the transposition operation, the whole PCA pan-sharpening process can be expressed by equation (2.1) using:

$$\mathbf{g} = \boldsymbol{\phi}_1{}^{\mathrm{T}}, \boldsymbol{Y}^{\mathrm{l}} = \boldsymbol{\phi}_1 \times \widetilde{\mathbf{X}}^{\mathrm{l}}\,. \tag{2.2}$$

GS and GSA: in the GS method, $\boldsymbol{Y}^{\mathrm{l}}$ is calculated by simply averaging the HS bands, while in its modified version, GSA, $\boldsymbol{Y}^{\mathrm{l}}$ is obtained through a linear transformation of the upsampled HS bands. The coefficients of the transformation are determined through the least squares method by establishing a linear relationship between a downsampled Pan image ($\widetilde{\boldsymbol{Y}}^{\mathrm{l}}$) and the original HS bands ($\mathbf{X}^{\mathrm{l}}$). This results in adaptively calculating the contribution of each spectral band in the fused image. The injection gain for the i-th band, in both methods, is defined by the ratio of the covariance between $\widetilde{\boldsymbol{x}}_i^{\mathrm{l}}$ and $\boldsymbol{Y}^{\mathrm{l}}$ over the variance of $\boldsymbol{Y}^{\mathrm{l}}$ (Loncan et al., 2015). In other words, if $\boldsymbol{W} \in \mathbb{R}^{1\times m_h}$ shows the weights vector, $\boldsymbol{Y}^{\mathrm{l}}$ and $\mathbf{g}$ in equation (2.1) are defined using:

$$\mathbf{g} = cov(\widetilde{\mathbf{X}}^{\mathrm{l}}, \boldsymbol{Y}^{\mathrm{l}})/var(\boldsymbol{Y}^{\mathrm{l}}), \boldsymbol{Y}^{\mathrm{l}} = \boldsymbol{W} \times \widetilde{\mathbf{X}}^{\mathrm{l}},$$

$$\boldsymbol{W} = \begin{cases} \left[\frac{1}{m_h}\ \frac{1}{m_h} \dots \frac{1}{m_h}\right], & \text{in GS} \\ argmin\ \left\|\widetilde{\boldsymbol{Y}}^{\mathrm{l}} - \boldsymbol{W} \times \mathbf{X}^{\mathrm{l}}\right\|, & \text{in GSA} \end{cases}. \tag{2.3}$$

2.2.1.2 MRA methods

In these methods, the spatial details are extracted from the high resolution image using multiresolution decompositions, such as Generalized Laplacian Pyramids (GLP) or different variations of Wavelet Transformation (WT). However, the WT-based methods cannot be simply adapted to HS pan-sharpening since the spatial resolution ratio between Pan and HS images is not usually an integer or the power of two (Loncan et al., 2015). Thus, $\boldsymbol{Y}^{\mathrm{l}}$ in these methods is obtained through a multiresolution decomposition of the Pan image. Furthermore, injection gains are not constant for all the pixels in the i-th band anymore; i.e. $\mathbf{g} \in \mathbb{R}^{m_h \times n_p}$.

SFIM: this method can be considered as the simplest method in the MRA group since it decomposes the Pan image only to one resolution level and using a simple low pass filter (LPF). $\boldsymbol{Y}^{\mathrm{l}}$ in this method is obtained by applying an LPF to the Pan image. The LPF is usually set as an averaging filter over a neighbourhood equivalent to the actual resolution of the low resolution spectral bands (Loncan et al., 2015). The injection gain for the i-th HS band is the ratio of the pixel values in that band over those of $\boldsymbol{Y}^{\mathrm{l}}$. If the symbol $*$ denotes the convolution operator and $\boldsymbol{k}_{\mathrm{LP}}$ shows the kernel used for filtering, we can write

$$\mathbf{g} = \frac{\widetilde{\mathbf{X}}^{\mathrm{l}}}{\boldsymbol{Y}^{\mathrm{h}} * \boldsymbol{k}_{\mathrm{LP}}}, \boldsymbol{Y}^{\mathrm{l}} = \boldsymbol{Y}^{\mathrm{h}} * \boldsymbol{k}_{\mathrm{LP}}. \tag{2.4}$$

GLP-HPM: in this method, the decomposition of the Pan image is performed in more than one resolution level. In the first resolution level, a Gaussian filter is applied to the Pan image to get a smoothed version of the Pan image and to obtain the next resolution levels, the Gaussian filter is applied to the smoothed image at each level. The Laplacian pyramid is then achieved by subtracting the smoothed images in the consecutive levels, which results in extracting the spatial details from the Pan image in different resolution levels. The details are then modulated, resulting in the multiplicative injection (B Aiazzi et al., 2006). If L shows the maximum level of decomposition, and $\boldsymbol{k}_{\mathrm{G}}$ represents the kernel of Gaussian filter, $\boldsymbol{Y}^{\mathrm{l}}$ and $\mathbf{g}$ are then defined through:

$$\mathbf{g} = \boldsymbol{\beta} \frac{\widetilde{\mathbf{X}}^{\mathrm{l}}}{\boldsymbol{Y}^{\mathrm{l}}}, \boldsymbol{Y}^{\mathrm{l}} = \boldsymbol{Y}^{\mathrm{h}} * \underbrace{\boldsymbol{k}_{\mathrm{G}} * \boldsymbol{k}_{\mathrm{G}} * \ldots * \boldsymbol{k}_{\mathrm{G}}}_{L \text{ times}}, \tag{2.5}$$

where $\boldsymbol{\beta}$ is constant for all the bands and is defined as the ratio between the average of the standard deviations of the HS bands and the standard deviation of $\boldsymbol{Y}^{\mathrm{l}}$.

2.2.1.3 UNB-Pansharp Method

This method cannot be grouped in CS or MRA groups, however, it can be considered as one of the methods with multiplicative sharpening model. In this method, a linear relationship between the Pan image and the HS bands is established, whose coefficients are determined through the least square technique. Thus, the best possible relationship between the spectral bands and the Pan image is obtained which results in adaptively including the contribution of each spectral band in the fused image. Due to the statistical approach in UNB-Pansharp method, the spectral relationship between the input spectral bands and the Pan band is estimated automatically (Y. Zhang, 2008; Y. Zhang & Mishra, 2014).

2.2.2 Pan-sharpening Methods Developed for HS Images

These methods work mainly based on two sensor observation models and a dimensionality reduction model. Each observation model relates the unknown high resolution HS image with one of the observed images, Pan or low resolution HS image. The dimensionality reduction model, on the other hand, tries to find a lower dimensional subspace in which the HS image "lives" (Loncan et al., 2015).

The observation models consider the spatial and spectral degradations between the $\mathbf{X}^{\mathrm{h}}$ and the observations, $\mathbf{X}^{\mathrm{l}}$ and $\boldsymbol{Y}^{\mathrm{h}}$, respectively. In other words, we can write (Yokoya et al., 2012)

$$\mathbf{X}^{\mathrm{l}} = \mathbf{X}^{\mathrm{h}}\mathbf{S} + \mathbf{N}_{\mathrm{X}},\ \boldsymbol{Y}^{\mathrm{h}} = \mathbf{R}\mathbf{X}^{\mathrm{h}} + \mathbf{N}_{\mathrm{Y}}, \tag{2.6}$$

where $\mathbf{S} \in \mathbb{R}^{n_p \times n_h}$ is the spatial degradation matrix which is in fact a product of a blurring matrix, representing the point spread function (PSF) of the sensor, by a downsampling matrix, and $\mathbf{R} \in \mathbb{R}^{1 \times m_h}$ is the spectral response matrix. The matrices $\mathbf{S}$ and $\mathbf{R}$ in fact describe the sensor specifications and are either considered to be known, such as in CNF method (Yokoya et al., 2012), or are estimated in the fusion procedure, such as in HySure method (Simoes et al., 2015). $\mathbf{N}_{\mathrm{X}}$ and $\mathbf{N}_{\mathrm{Y}}$ are the related observation noises.

The dimensionality reduction model helps to make algorithms faster and more accurate since the number of variables to be estimated is reduced. If $\mathbf{E} \in \mathbb{R}^{m_h \times d}$ denotes the transformation matrix containing d bases related to the new subspace ($d < m$), and $\mathbf{W} \in \mathbb{R}^{d \times n_p}$ represents the corresponding coefficients, $\mathbf{X}^{\mathrm{h}}$ can be converted to the lower dimensional subspace through (Loncan et al., 2015):

$$\mathbf{X}^{\mathrm{h}} = \mathbf{EW}. \tag{2.7}$$

Equation (2.7) is, in fact, the main fusion model considered in the HS pan-sharpening methods, for which the matrices $\mathbf{E}$ and $\mathbf{W}$ must be estimated. To obtain $\mathbf{E}$ and $\mathbf{W}$, CNMF method uses matrix factorization technique, while in HySure method a Bayesian approach is applied. Before reviewing these two methods, here we show that equation (2.7), with some approximations, can be reformed in a way that assimilates the GFM in equation (2.1), no matter how the matrices $\mathbf{E}$ and $\mathbf{W}$ are estimated. To do so, we need to obtain a relationship which relates $\mathbf{X}^{\mathrm{h}}$ to the observed images, $\mathbf{X}^{\mathrm{l}}$ and $\boldsymbol{Y}^{\mathrm{h}}$. In Section 2.2.3 we discuss how this reformation, helps for adaptive band selection for the purpose of pan-sharpening. By substituting $\mathbf{X}^{\mathrm{h}}$ from equation (2.7) in equations (2.6) and by ignoring the observation noises, we have

$$\mathbf{X}^{\mathrm{l}} \approx \mathbf{EWS}, \tag{2.8}$$

$$\boldsymbol{Y}^{\mathrm{h}} \approx \mathbf{REW}. \tag{2.9}$$

Now, we can obtain $\mathbf{E}$ and $\mathbf{W}$ based on $\mathbf{X}^{\mathrm{l}}$ and $\boldsymbol{Y}^{\mathrm{h}}$, and by substituting them in equation (2.7) we get

$$\mathbf{X}^{\mathrm{h}} = \mathbf{X}^{\mathrm{l}}\mathbf{S}^{-1}\mathbf{W}^{-\mathbf{1}}\mathbf{E}^{-\mathbf{1}}\mathbf{R}^{-1}\boldsymbol{Y}^{\mathrm{h}}. \tag{2.10}$$

The matrix $\mathbf{S}^{-\mathbf{1}}$ can be considered as an operator, which acts in an opposite way to the matrix $\mathbf{S}$, i.e. it gives an upsampled version of the image by which it multiplies; thus, we can write $\widetilde{\mathbf{X}}^{\mathrm{l}} = \mathbf{X}^{\mathrm{l}}\mathbf{S}^{-\mathbf{1}}$, which gives

$$\begin{aligned}\mathbf{X}^{\mathrm{h}} = \widetilde{\mathbf{X}}^{\mathrm{l}}\mathbf{W}^{-\mathbf{1}}\mathbf{E}^{-\mathbf{1}}\mathbf{R}^{-1}\boldsymbol{Y}^{\mathrm{h}} &= \widetilde{\mathbf{X}}^{\mathrm{l}}(\mathbf{REW})^{-\mathbf{1}}(\boldsymbol{Y}^{\mathrm{h}} - \mathbf{REW} + \mathbf{REW}) \\ &= \widetilde{\mathbf{X}}^{\mathrm{l}} + \widetilde{\mathbf{X}}^{\mathrm{l}}(\mathbf{REW})^{-\mathbf{1}}(\boldsymbol{Y}^{\mathrm{h}} - \mathbf{REW}).\end{aligned} \tag{2.11}$$

Finally, substituting $\mathbf{EW}$ from equation (2.8) in equation (2.11) gives

$$\begin{aligned}\mathbf{X}^{\mathrm{h}} = \widetilde{\mathbf{X}}^{\mathrm{l}} + \widetilde{\mathbf{X}}^{\mathrm{l}}\left(\mathbf{R}\mathbf{X}^{\mathrm{l}}\mathbf{S}^{-\mathbf{1}}\right)^{-\mathbf{1}}\left(\boldsymbol{Y}^{\mathrm{h}} - \mathbf{R}\mathbf{X}^{\mathrm{l}}\mathbf{S}^{-\mathbf{1}}\right) &= \widetilde{\mathbf{X}}^{\mathrm{l}} + \widetilde{\mathbf{X}}^{\mathrm{l}}\left(\mathbf{R}\widetilde{\mathbf{X}}^{\mathrm{l}}\right)^{-\mathbf{1}}\left(\boldsymbol{Y}^{\mathrm{h}} - \mathbf{R}\widetilde{\mathbf{X}}^{\mathrm{l}}\right) \\ &= \widetilde{\mathbf{X}}^{\mathrm{l}} + \mathbf{R}^{-\mathbf{1}}\left(\boldsymbol{Y}^{\mathrm{h}} - \mathbf{R}\widetilde{\mathbf{X}}^{\mathrm{l}}\right).\end{aligned} \tag{2.12}$$

Now, comparing equation (2.12) with equation (2.1), we can set the injection gain and the low resolution version of the Pan image as:

$$\mathbf{g} = \mathbf{R}^{-\mathbf{1}}, \boldsymbol{Y}^{\mathrm{l}} = \mathbf{R}\widetilde{\mathbf{X}}^{\mathrm{l}}. \tag{2.13}$$

Thus, as can be seen, the procedure of CNMF and HySure methods can also be described by the model in equation (2.1). Although these methods do not directly use equation (2.12), and they rather estimate the matrices $\mathbf{E}$ and $\mathbf{W}$ to obtain the fused image, it is shown that neglecting the observation noises, their procedure mathematically can be approximated by equation (2.12). In the following sections, we briefly review how these methods estimate $\mathbf{E}$ and $\mathbf{W}$.

2.2.2.1 CNMF Method

In this method, equation (2.7) is considered as a linear spectral mixture model, which is commonly used for unmixing the hyperspectral images. With such a perspective, $\mathbf{E}$ is considered as the endmembers' spectra matrix and $\mathbf{W}$ is the abundance matrix. In fact, CNMF pan-sharpening alternately unmixes both observed HS and Pan images into endmember and abundance matrices, using the nonnegative matrix factorization. Sensor observation models in equation (2.6) are used for the initialization step in each HS/Pan unmixing procedure. Similar to the mixture model of $\mathbf{X}^{\mathrm{h}}$, two other models can be written for each observed image as (Yokoya et al., 2012):

$$\mathbf{X}^{\mathrm{l}} \approx \mathbf{E}\mathbf{W}_{\mathrm{S}}, \tag{2.14}$$

$$\boldsymbol{Y}^{\mathrm{h}} \approx \mathbf{E}_{\boldsymbol{R}}\mathbf{W}. \tag{2.15}$$

Where $\mathbf{W}_{\mathrm{S}}$ and $\mathbf{E}_{\boldsymbol{R}}$ are the endmember and the abundances matrices degraded in spatial and spectral domains, respectively. In other words (Yokoya et al., 2012):

$$\mathbf{W}_{\mathrm{S}} \approx \mathbf{W}\mathbf{S}, \tag{2.16}$$

$$\mathbf{E}_{\boldsymbol{R}} \approx \mathbf{R}\mathbf{E}. \tag{2.17}$$

In summary, the CNMF procedure contains three main steps (Yokoya et al., 2012): 1) unmixing of $\mathbf{X}^{\mathrm{l}}$ data, for which the matrix $\mathbf{E}$ is initialized using one of the unmixing algorithms, e.g. the vertex component analysis (VCA), and $\mathbf{E}$ and $\mathbf{W}_{\mathrm{S}}$ are alternately optimized by minimizing $\left\|\mathbf{X}^{\mathrm{l}} - \mathbf{E}\mathbf{W}_{\mathrm{S}}\right\|_{\mathrm{F}}^{2}$, where $\|.\|_{\mathrm{F}}^{2}$ denotes the Frobenius norm. 2) Next, unmixing of $\boldsymbol{Y}^{\mathrm{h}}$ data, where based on equation (2.17), $\mathbf{E}_{\boldsymbol{R}}$ is set to $\mathbf{R}\mathbf{E}$ and $\mathbf{W}$ is initialized by upsampling $\mathbf{W}_{\mathrm{S}}$ obtained in the previous step, and then is updated by minimizing $\left\|\boldsymbol{Y}^{\mathrm{h}}-\mathbf{E}_{\boldsymbol{R}}\mathbf{W}\right\|_{\mathrm{F}}^{2}$. 3) Finally, the high spatial resolution HS image is obtained by multiplying $\mathbf{E}$ and $\mathbf{W}$ based on equation (2.7).

2.2.2.2 HySure Method

This method works under a Bayesian framework and formulates the fusion problem as a convex program. In addition, the relative spectral and spatial responses of the sensors, which reflect in the matrices $\mathbf{R}$ and $\mathbf{S}$, are formulated as another convex quadratic program and are estimated from the observed images. The matrix $\mathbf{E}$ is estimated using one of the unmixing algorithm, e.g. VCA. For estimating the matrix $\mathbf{W}$, an optimization problem is defined which imposes that the estimated high resolution HS image should fit the observation models in equation (2.6). By substituting $\mathbf{X}^{\mathrm{h}}$ from equation (2.7) in equations (2.6), the optimization problem tackled in this method is formulated as (Simoes et al., 2015)

$$minimize_{\mathbf{W}}\ \frac{1}{2}\left\|\mathbf{X}^{\mathrm{l}}-\mathbf{EWS}\right\|_{\mathrm{F}}^{2}+\frac{\lambda_{\mathrm{m}}}{2}\left\|\boldsymbol{Y}^{\mathrm{h}}-\mathbf{REW}\right\|_{\mathrm{F}}^{2}+\lambda_{\phi}\phi(\mathbf{W}), \tag{2.18}$$

where the parameters λ_{m} and λ_{φ} are used to control the contribution of different terms. $\lambda_{\varphi}\varphi(\mathbf{W})$ is the regularizer, which is defined as a form of Vector Total Variation (Simoes et al., 2015).

To estimate the matrices $\mathbf{R}$ and $\mathbf{S}$, another optimization problem should be solved. Ignoring the noise parameters in equations (2.6), we can write $\mathbf{RX}^{\mathrm{l}}=\boldsymbol{Y}^{\mathrm{h}}\mathbf{S}$, which is used along with two other regularizers, $\lambda_{\mathrm{R}}\phi_{\mathrm{R}}(\mathbf{R})$ and $\lambda_{\mathrm{S}}\phi_{\mathrm{S}}(\mathbf{S})$, to form the following quadratic objective function (Simoes et al., 2015):

$$minimize_{\mathbf{S},\mathbf{R}}\ \left\|\mathbf{RX}^{\mathrm{l}}-\boldsymbol{Y}^{\mathrm{h}}\mathbf{S}\right\|^{2}+\lambda_{\mathrm{R}}\phi_{\mathrm{R}}(\mathbf{R})+\lambda_{\mathrm{S}}\phi_{\mathrm{S}}(\mathbf{S}). \tag{2.19}$$

2.2.3 Adaptive Band Selection

The effect of inappropriate band selection for the pan-sharpening procedure shows itself usually as the low spectral quality in the final fused image. This effect can be traced in the fusion model used in each method. In the previous sections, it was shown that the reviewed pan-sharpening methods, exactly or approximately, follow the GFM in equation (2.1). Moreover, equation (2.1) relates the fused image with the observed Pan and HS images directly, and thus, provides a convenient way to study the effect of any spectral discrepancies between Pan and HS images on the fused image.

Based on equation (2.1), the term that causes the spectral distortion problem is $(\boldsymbol{Y}^{\mathrm{h}} - \boldsymbol{Y}^{\mathrm{l}})$, i.e. the difference between the observed Pan image and its low resolution version. In other words, the higher the spectral discrepancies between $\boldsymbol{Y}^{\mathrm{h}}$ and $\boldsymbol{Y}^{\mathrm{l}}$, the higher the spectral distortion is expected to occur in the fused images. Therefore, to prevent any spectral incompatibility, the Pan band and its low resolution version should be spectrally similar. This spectral similarity is already guaranteed in the methods where $\boldsymbol{Y}^{\mathrm{l}}$ is obtained only from the Pan band, as in MRA methods (see Section 2.2.1.2). However, what band selection affects in these methods is the quality of the spatial details modulated through the injection gain.

Here, we first focus on those methods where $\boldsymbol{Y}^{\mathrm{l}}$ is obtained from the low resolution HS bands (as in CS methods, UNB-Pansharp, CNMF, and HySure methods), and we establish the adaptive band selection mainly for these methods. We then will show that the suggested band selection approach also helps in MRA methods to get better spatial qualities when having noisy bands in the HS data.

In CS methods as well as UNB-Pansharp, CNMF, and HySure methods, $\boldsymbol{Y}^{\mathrm{l}}$ is calculated from the low resolution HS. Therefore, to prevent any spectral incompatibility, the Pan band should be spectrally dependent on the HS bands which are selected to generate $\boldsymbol{Y}^{\mathrm{l}}$. This makes it clear why in practice, the HS bands which fall within the spectral range of the Pan band are selected for the pan-sharpening. In fact, the same spectral range covered by Pan and HS bands guarantees the high spectral dependency between Pan and HS bands. This motivates the idea of looking for other spectral bands outside of the Pan spectral range, which may be spectrally similar to the bands within the spectral range of the Pan band, knowing that the HS bands are usually highly correlated due to the very close spectral ranges they cover. Since correlation is one of the statistical tools to measure the amount of dependency of two variables, we then establish our proposed band selection rule based on the correlation. This states that any HS band that is highly correlated with the Pan band can be a good candidate to be incorporated in the pan-sharpening procedure.

Let $\mathrm{X}_{\mathrm{S}}^{\mathrm{l}} = \left\{ \boldsymbol{x}_i^{\mathrm{l}} \,\middle|\, S_{\boldsymbol{x}_i^{\mathrm{l}}} \subset S_{\boldsymbol{Y}^{\mathrm{h}}} \right\}$ represent the subset of HS bands whose spectral ranges, $S_{\boldsymbol{x}_i^{\mathrm{l}}}$, fall within the spectral range of the Pan band, $S_{\boldsymbol{Y}^{\mathrm{h}}}$, and $\widetilde{\mathbf{X}}_{\mathrm{S}}^{\mathrm{l}} \in \mathbb{R}^{m_h{'} \times n_p}$ denote the matrix containing the upsampled bands with a total number of $m_h{'}$. If F denotes the fusion function representing the GFM, we can rewrite equation (2.1) as:

$$\mathbf{X}^{\mathrm{h}} = F\left(\widetilde{\mathbf{X}}_{\mathrm{S}}^{\mathrm{l}}, \boldsymbol{Y}^{\mathrm{h}}, \boldsymbol{Y}^{\mathrm{l}}\right). \tag{2.20}$$

We then propose to create a new subset of bands, $\mathrm{X}_{\mathrm{S}'}^{\mathrm{l}} = \left\{ \boldsymbol{x}_i^{\mathrm{l}} \middle| \, \mathrm{Corr}\left(\boldsymbol{x}_i^{\mathrm{l}}, \boldsymbol{Y}^{\mathrm{h}}\right) \geq \tau \right\}$, where τ is a threshold to select highly correlated bands. If $\widetilde{\mathbf{X}}_{\mathrm{S}'}^{\mathrm{l}} \in \mathbb{R}^{m_h{''} \times n_p}$ represents the matrix containing the upsampled bands in $\mathrm{X}_{\mathrm{S}}^{\mathrm{l}}$, with a total number of m_h", equation (2.20) can be rewritten as:

$$\mathbf{X}^{\mathrm{h}} = \mathrm{F}\left(\widetilde{\mathbf{X}}_{\mathrm{S}',}^{\mathrm{l}}, \boldsymbol{Y}^{\mathrm{h}}, \boldsymbol{Y}^{\mathrm{l}}\right) \tag{2.21}$$

Observations also confirm that the Pan band is correlated not only with the bands in $\mathrm{X}_{\mathrm{S}}^{\mathrm{l}}$, but with other bands beyond the Pan spectral range. These observations will be presented in Section 2.3.2. It is noteworthy that we use linear correlation since the GFM is a linear model and all the methods studied in this chapter fit into that linear model.

In MRA methods, however, $\boldsymbol{Y}^{\mathrm{l}}$ is obtained using only the Pan band and thus, there is no spectral discrepancy between $\boldsymbol{Y}^{\mathrm{h}}$ and $\boldsymbol{Y}^{\mathrm{l}}$. This explains why in the literature MRA methods are reported to have low spectral distortion (Loncan et al., 2015). Therefore, by selecting even all the HS bands, no huge spectral changes are expected, since the HS bands do not play any role in the term ($\boldsymbol{Y}^{\mathrm{h}} - \boldsymbol{Y}^{\mathrm{l}}$). However, there is another term in MRA method, which relates the Pan and HS bands, and that is the injection gain. The injection gain, based on equations (2.4) and (2.5), is defined as the ratio of the low resolution HS bands to the low-pass filtered Pan image. This ratio modulates the spatial details to the fused bands. Thus, if the HS dataset contains noisy bands (which is always the case in reality), selecting all bands for the MRA methods results in injecting the noise from the noisy HS bands to the fused bands. However, with selecting bands that are highly correlated with Pan, the possibility of selecting those noisy bands decreases since they show low correlation with Pan band (this will be shown in Section 2.3.2). Therefore, for the MRA methods, it is expected that the suggested band selection approach provides better spatial quality compared to selecting all HS bands, while increasing the number of bands for the pan-sharpening.

2.3 Results and Discussion

Summarizing the experiment, after selecting the appropriate datasets, the correlation values between the Pan image and the HS bands were calculated which provided the criteria to select the highly correlated HS bands with the Pan image as the suitable bands that can be fused. Then, the selected HS bands were fused with Pan image using the methods described in Sections 2.2.1 and 2.2.2. Finally, the results are evaluated. The details of these steps are explained in the following subsections.

2.3.1 Datasets

The images we selected for this experiment are from EO-1 satellite. The reason why we selected EO-1 satellite is that it provides simultaneous Pan and HS images; however, the proposed band selection approach can be used to fuse any other HS images (even the airborne ones) with any Pan images taken simultaneously. The Pan image of the EO-1 data set is captured by the Advanced Land Imager (ALI) sensor and the HS image is captured by the Hyperion sensor. The ALI sensor provides Pan band with 10 meter resolution and with spectral coverage from 480 nm to 690 nm. The Hyperion sensor provides a 30 meter resolution HS image with 242 spectral bands sampled by 10 nm intervals, which cover the spectrum from 400 nm to 2500 nm, or from visible to near-infrared (NIR) range (Griffin et al., 2005). From 242 HS bands, 44 bands are not calibrated (B_1 to B_7, B_{58} to B_{76}, and B_{225} to B_{242}) and hence are removed from the original images before the processing. From the remaining 198 bands (B_8 to B_{57} and B_{77} to B_{224}), only 20 bands fall within the spectral range of the Pan band, i.e., $\mathrm{X}_{\mathrm{S}}^{\mathrm{L}} = \{\boldsymbol{x}_{14}^{\mathrm{l}}, \boldsymbol{x}_{15}^{\mathrm{l}}, \dots, \boldsymbol{x}_{33}^{\mathrm{l}}\}$, which belongs mainly to the visible range of the electromagnetic spectrum, 488 nm to 681 nm in average wavelength.

For the experimental tests, three different datasets were used, which cover some parts of Calgary (Canada), Tehran (Iran), and Clovis (USA) cities including a variety of landscapes such as urban, rural, and mountainous areas with different crops, vegetations, soils, and water body coverages. Table 2.1 summarizes the specification of these datasets. It is noteworthy that we have not limited our investigation to a small area of the images, i.e. the dimensions of the images are set to be close to the real-world applications, which makes the study of the methods' behaviour more realistic. All datasets were selected from level 1T (L1T), which means that radiometric and systematic geometric corrections using ground control points had previously been applied. Nevertheless, the Pan and HS images in the Calgary and the Tehran datasets were not well aligned, and thus we registered them before the fusion step, but for the Clovis dataset, the images were georeferenced with high accuracy and no further registration was required.

Table 2. 1 Characteristics of the datasets.

Dataset	**Dimensions**	**Acquisition date**	**Landscape**
Calgary, Alberta, Canada	Pan: 3210 x 1734 HS: 1070 x 578	11 Jul. 2003	a mixed urban/rural scene with different kinds of crops
Tehran, Iran	Pan: 3600 x 1650 HS: 1200 x 550	26 Sep. 2004	a mixed urban /rural/ mountainous scene with different kinds of crops
Clovis, New Mexico, USA	Pan: 2997 x 1233 HS: 999 x 411	4 May. 2003	a mixed urban/rural scene with different kinds of crops

2.3.2 Adaptive Band Selection Using Correlation Criteria

As explained in Section 2.2.3, we use correlation to seek for any possible dependencies between Pan and HS bands, which are beyond the Pan spectral range. For this purpose, correlation values between Pan and all HS bands are calculated for all the datasets and are presented in Figure 2.1. The areas with no correlation values are related to the non-calibrated bands that were removed from the original images before the processing, and the areas with very low correlation values are mainly related to the noisy bands.

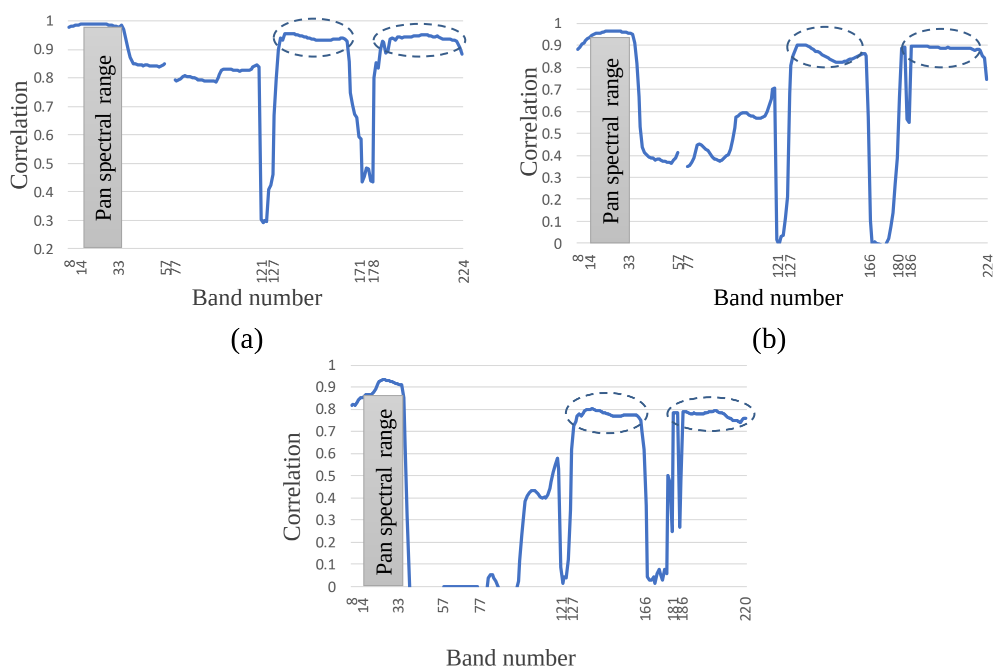


Figure 2. 1 Correlation between Pan and HS bands in (a) Calgary, (b) Tehran, and (c) Clovis datasets. The highlighted areas show the bands which are beyond the Pan spectral range but are highly correlated with the Pan band.

As can be seen in Figure 2.1, in all three datasets, the spectral bands which fall within the Pan spectral range (B_{14} to B_{33}) show the highest correlation values with the Pan band, as was expected. However, the interesting observation is that there are other spectral bands, specified with dashed circles in the plots, that are also highly correlated with the Pan band, although they belong to the spectral ranges beyond the Pan spectral range. Moreover, comparing the correlation plots of the different datasets shows that almost the same spectral bands, ranging approximately from 1400 nm to 1800 nm, and 2000 nm to 2400 nm, show high correlation with the Pan band. In addition, as can be observed in the correlation plots of Figure 2.1, the neighbouring spectral bands have almost the same correlation values.

Selecting an optimum correlation threshold (τ), which results in selecting appropriate bands for the pan-sharpening procedure, is in fact a trade-off between the number of bands and the amount of spectral distortion that may appear in the fusion results. The optimum threshold for each dataset is selected based on the calculated correlation values, so that an acceptable level of spectral distortion is maintained. For this purpose, different threshold candidates were selected for each dataset; for example, in the Calgary dataset, the spectral bands can be roughly categorized into 6 classes, B_8-B_{36}, B_{37}-B_{57}, B_{77}-B_{100}, B_{100}-B_{120}, B_{128}- B_{162}, B_{189}- B_{219}, with average correlation values as 0.99, 0.85, 0.80, 0.83, 0.95, and 0.94, respectively. Each of these values can be considered as a candidate threshold that introduces different sets of spectral bands from different areas of the spectrum to be fed into the pan-sharpening procedure. It is noteworthy that among these 6 classes, we exclude the first class from the threshold selection procedure, because it is mainly related to the Pan spectral range and we want to find other appropriate bands beyond the Pan spectral range. Then, for each candidate threshold, the related bands were fed to the pan-sharpening procedure and the spectral signatures were obtained from the fused bands. The threshold value which results in reconstructing the most similar spectral signature to the original one while including more bands, is then selected as the optimum one.

Table 2.2 reports the optimum thresholds and the bands that were selected for the pan-sharpening procedure in each dataset. As can be seen, we mainly exclude B_{38} to B_{127} (730-1800nm) from the pan-sharpening procedure in all three datasets.

Table 2. 2 Selected bands for pan-sharpening procedure in each dataset.

Dataset	**Threshold (τ)**	**Bands highly correlated with Pan**	**Total number of bands**
Calgary	0.94	B_8-B_{36}, B_{128}-B_{162}, B_{189}-B_{219}	95
Tehran	0.80	B_8-B_{37}, B_{128}-B_{165}, B_{179}-B_{224}	114
Clovis	0.75	B_8-B_{36}, B_{128}-B_{165}, B_{182}-B_{221}	107

2.3.3 Comparison of Pan-sharpening Results

The pan-sharpening methods, reviewed in Sections 2.2.1 and 2.2.2, have been applied to the three datasets summarized in Table 2.1. Moreover, in order to compare the adaptive band selection with the other band selection scenarios in the literature, three fusion results were generated from each pan-sharpening method using: 1) the HS bands within the Pan spectral range (all the bands in X_S^l subset), 2) the HS bands that are highly correlated with Pan (all the bands in X_S^l, subset), and 3) all the HS bands. Table 2.3

summarizes the band numbers and the total number of bands for each band selection scenario.

Table 2. 3 Different band selection scenarios in this experiment.

Band selection scenarios	**Bands**	**Total number of bands**
Bands within the Pan spectral range	B_{14}-B_{33}	20
Bands highly correlated with Pan	As reported in Table 2.2.	
All bands (excluding non-calibrated bands)	B_8-B_{57}, B_{77}-B_{224}	198

2.3.3.1 Visual Analysis

Figures 2.2 shows three small areas cropped from the fused images of the mentioned datasets and obtained from the considered methods as well as the original HS and Pan images as the references. For visualization, the bands to form the colour composition are selected from the visible range (R= B_{33}, G= B_{23}, B= B_{14}).

A visual analysis of the results in Figure 2.2 shows that different fused images contain different amounts of spectral distortion depending on what method or which HS bands are used. For PCA, GS, and CNMF methods comparing the results from left to right in all datasets, it can be observed that as expected the minimum spectral distortion belongs to the case where only 20 bands in the Pan spectral range are used (B_{14}-B_{33}). On the other hand, using all bands yields the highest amount of spectral distortion. Using bands highly correlated with Pan, although it results in slightly higher spectral distortion compared to when 20 bands are used, still gives more acceptable results compared to using all bands. In contrast with PCA and GS methods, for the GSA and UNB-Pansharp methods, no meaningful differences in terms of spectral fidelity can be observed among the fused images obtained from using different sets of bands. This shows that these methods are less sensitive to the bands fed to them. This is also the case for GLP-HPM and SFIM methods, however, belonging to MRA methods, their results are not as sharp as other results. In fact, they are not successful in injecting high-quality spatial details that exist in the Pan image through the term $\boldsymbol{Y}^{\mathrm{h}} - \boldsymbol{Y}^{\mathrm{l}}$ (equation 2.1), in which $\boldsymbol{Y}^{\mathrm{l}}$ is obtained by applying low pass filters to the Pan image. In addition, some colour bleeding can be observed along the edges of different objects in the image. This shows that these methods are very sensitive to even small misregistration which is inevitable while registering Pan and HS bands. As mentioned before, what band selection affects in the MRA methods is the quality of the spatial details modulated through the injection gain. Figure 2.3 is provided to visually investigate this issue. The images are the results of SFIM method and are cropped from Tehran dataset with the colour composition of R= B_{165}, G= B_{155}, B= B_{135}, which belong to the range outside the Pan spectral range. In the original HS image of Figure 2.3, some noises in the form of pink and cyan strips are observed which have been transferred into the fused image obtained using all bands but are removed in the fused image obtained from bands highly correlated with Pan.

For HySure method, as it can be seen especially in the Calgary and Clovis datasets, using bands highly correlated with Pan produces less spectral distortion compared to the case where all bands are used, and it has not introduced much distortion compared to when 20 bands are used. However, our observations show that along the edges some spatial artefacts as well as colour bleeding exist, and these effects intensify when all bands are used.

Calgary Dataset
Original HS
Pan
PCA
GS
GSA
UNB-Pansharp
SFIM
GLP-HPM
CNMF
HySure
Tehran Dataset
Original HS
Pan
PCA
GS
GSA
UNB-Pansharp
SFIM
GLP-HPM
CNMF
HySure
Clovis Dataset
Original HS
Pan
PCA
GS

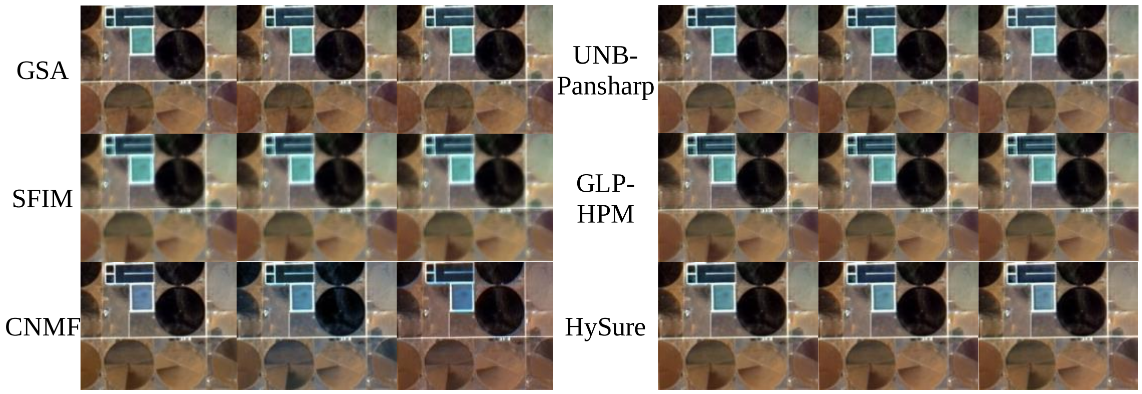


Figure 2. 2 Cropped areas of original HS image, and the fused results by different pan-sharpening methods in different datasets (colour composition: R= B33, G= B23, B= B14 selected from the visible range). Three fused images are obtained through each method using, from left to right, the bands that share common spectral range with Pan, the bands that are highly correlated with Pan, and all HS bands.

Figure 2. 3 Comparing the original HS image containing noisy bands (left), and the SFIM results using bands highly correlated with Pan (middle), and all HS bands (right). The colour composition is R= B165, G= B155, B= B135. The highlighted areas show the strip noises in the original HS image, which are directly transferred to the right image but disappear from the left image.

2.3.3.2 Quality Indices

To evaluate the quality of the fused images, some quality indices have been defined and used in the literature, which mainly calculates the similarity between the fused image and the original MS/HS image. Since the reference image, the high resolution HS image, is not available in the case of pan-sharpening, Wald's protocol (Loncan et al., 2015) is followed, which suggests obtaining the degraded fused image in a reduced resolution level (which we show with $\tilde{\mathbf{X}}^{\mathrm{h}}$) and to compare it with the original low resolution HS image as the reference. For this experiment, we used Spectral Angle Mapper (SAM), Erreur Relative Globale Adimensionnelle de Synth`ese (ERGAS) as the most widely used quality indices (Loncan et al., 2015), and Quality with No Reference (QNR; Alparone et al., 2008). SAM and ERGAS are mainly considered as the spectral quality indices. QNR, however, does not require a reference image, and consists of both spectral term, assessed at reduced resolution level, and spatial term assessed at full resolution level. Here, we first briefly review each quality index and then we provide the numerical results achieved from our experiment.

- ERGAS is the relative global dimensional synthesis error and is independent of resolution and number of bands. If d shows the resolution ratio between the original HS and Pan images, ERGAS is defined as (Loncan et al., 2015)

  $$\mathrm{ERGAS} = 100d\sqrt{\frac{1}{n_p}\sum_{i=1}^{n_p}(\frac{\mathrm{RMSE}(i)}{\mu_i})^2}, \tag{2.22}$$

  where μ_i is the mean value of the i-th HS band, and $\mathrm{RMSE}(i)$ is the root mean square error calculated using the i-th band of the HS and fused images as $\mathrm{RMSE}(i) = \frac{\left\|\tilde{x}_i^{\mathrm{h}} - x_i^{\mathrm{l}}\right\|}{\sqrt{n_h}}$, in which $\|.\|$ denotes the ℓ_2 norm. The ideal value of ERGAS is zero.

- SAM is the angle between the spectral vector of each fused pixel and that of each original HS pixel (Loncan et al., 2015),

$$\mathrm{SAM}(F, B) = \cos^{-1}\frac{\langle \tilde{\boldsymbol{x}}_i^{\mathrm{h}}, \boldsymbol{x}_i^{\mathrm{l}} \rangle}{\|\boldsymbol{x}_i^{\mathrm{h}}\|\|\boldsymbol{x}_i^{\mathrm{l}}\|}, \tag{2.23}$$

where $\langle ., . \rangle$ denotes the inner product between two spectral vectors. SAM is calculated for each pixel and the final value is obtained by averaging SAM over all pixels. The ideal value of SAM is zero.

- QNR is, in fact, a jointly spectral and spatial quality index consisting of one spectral distortion index (D_λ) and one spatial distortion index (D_S) and is defined as (Alparone et al., 2008):

$$\mathrm{QNR} = (1 - D_\lambda)^\alpha . (1 - D_\mathrm{S})^\beta, \tag{2.24}$$

where α and β are real-valued exponent that define the contribution of spectral and spatial distortions to the overall quality. In this experiment, we set both to 1. The optimal value of QNR is one and can be obtained if the spectral and spatial distortions are both zero. The spectral and spatial distortion indices are then defined as (Alparone et al., 2008):

$$D_\lambda = \sqrt[p]{\frac{1}{n_p(n_p-1)}\sum_{i=1}^{n_p}\sum_{\substack{j=1 \\ j\neq 1}}^{n_p}\left|Q\left(\boldsymbol{x}_i^{\mathrm{l}}, \boldsymbol{x}_j^{\mathrm{l}}\right) - Q\left(\tilde{\boldsymbol{x}}_i^{\mathrm{h}}, \tilde{\boldsymbol{x}}_j^{\mathrm{h}}\right)\right|^p}, \tag{2.25}$$

$$D_\mathrm{S} = \sqrt[q]{\frac{1}{n_p}\sum_{i=1}^{n_p}\left|Q\left(\boldsymbol{x}_i^{\mathrm{l}}, \tilde{\boldsymbol{Y}}^{\mathrm{h}}\right) - Q\left(\tilde{\boldsymbol{x}}_i^{\mathrm{h}}, \boldsymbol{Y}^{\mathrm{h}}\right)\right|^q}, \tag{2.26}$$

in which p and q are positive integer exponents chosen to emphasize large spectral and spatial differences. In this experiment, we set both to 1. $\tilde{\boldsymbol{Y}}^{\mathrm{h}}$ is the spatially degraded version of the Pan image which is obtained by filtering the

Pan image with a Gaussian-shaped kernel approximating the modulation transfer function (MTF) of the HS sensor, followed by downsampling. Q is the Universal Image Quality Index (UIQI) and is calculated using the standard deviation, mean values, and the covariance between two vectors of interest $\boldsymbol{a}$ and $\boldsymbol{b}$ as (Alparone et al., 2008):

$$Q(\boldsymbol{a},\boldsymbol{b}) = \frac{\sigma_{\boldsymbol{ab}}}{\sigma_{\boldsymbol{a}} \cdot \sigma_{\boldsymbol{b}}} \cdot \frac{2\bar{a}}{[(\bar{a})^2 + (\bar{b})^2]} \cdot \frac{2\sigma_{\boldsymbol{a}} \cdot \sigma_{\boldsymbol{b}}}{(\sigma_{\boldsymbol{a}}{}^2 + \sigma_{\boldsymbol{b}}{}^2)}. \tag{2.27}$$

The mentioned quality indices have been measured on the fused results and are reported in Tables 2.4, 2.5, and 2.6 for different datasets. To obtain these values in each case, all the bands used for fusion are compared to the corresponding bands in the original HS image. Comparing the values of ERGAS, SAM, and D_λ for all the methods shows that the spectral quality of the fused images obtained from the bands highly correlated with Pan are very close to (and in some cases even better than) those reported for the fused images obtained from bands within Pan spectral range. Also, the spectral quality of the fused images obtained using the bands highly correlated with Pan are superior to that of the fused images obtained from all bands.

Table 2. 4 Measured quality indices for the Calgary dataset.

Fusion method	Selected bands for fusion	ERGAS	SAM	D_λ	D_S	QNR
PCA	Bands in Pan spectral range	6.51	4.53	0.07	0.10	0.83
	Highly correlated bands with Pan	7.09	4.35	0.21	0.21	0.63
	All bands	19.63	15.45	0.62	0.36	0.24
GS	Bands in Pan spectral range	6.52	4.82	0.08	0.10	0.83
	Highly correlated bands with Pan	8.18	5.96	0.21	0.24	0.60
	All bands	19.30	14.56	0.58	0.35	0.27
GSA	Bands in Pan spectral range	6.18	4.37	0.07	0.10	0.83
	Highly correlated bands with Pan	6.23	3.95	0.15	0.13	0.74
	All bands	9.43	4.29	0.20	0.17	0.67
UNB-Pansharp	Bands in Pan spectral range	6.04	3.39	0.04	0.10	0.86
	Highly correlated bands with Pan	7.84	4.48	0.12	0.14	0.75
	All bands	11.92	5.81	0.26	0.25	0.55
SFIM	Bands in Pan spectral range	2.92	2.26	0.04	0.09	0.88
	Highly correlated bands with Pan	3.55	2.53	0.12	0.09	0.81
	All bands	7.87	3.38	0.13	0.09	0.80
GLP-HPM	Bands in Pan spectral range	3.45	1.87	0.03	0.07	0.90
	Highly correlated bands with Pan	3.96	2.18	0.07	0.06	0.87
	All bands	8.20	3.22	0.08	0.07	0.85
CNMF	Bands in Pan spectral range	6.79	5.96	0.07	0.11	0.83
	Highly correlated bands with Pan	8.94	9.97	0.17	0.13	0.72
	All bands	18.20	18.86	0.24	0.18	0.62
HySure	Bands in Pan spectral range	5.32	4.67	0.18	0.15	0.69
	Highly correlated bands with Pan	7.94	6.57	0.19	0.20	0.65
	All bands	21.67	7.61	0.25	0.25	0.56

Table 2. 5 Measured quality indices for the Tehran dataset.

Fusion method	Selected bands for fusion	ERGAS	SAM	D_λ	D_S	QNR
PCA	Bands in Pan spectral range	2.85	2.87	0.06	0.02	0.93
	Highly correlated bands with Pan	3.12	3.71	0.11	0.04	0.86
	All bands	8.29	10.25	0.42	0.12	0.51
GS	Bands in Pan spectral range	2.85	2.88	0.08	0.10	0.83
	Highly correlated bands with Pan	3.70	4.01	0.21	0.24	0.60
	All bands	7.93	9.85	0.58	0.35	0.27
GSA	Bands in Pan spectral range	2.88	4.37	0.07	0.10	0.83
	Highly correlated bands with Pan	3.16	3.95	0.15	0.13	0.74
	All bands	5.93	4.29	0.20	0.17	0.67
UNB-Pansharp	Bands in Pan spectral range	2.81	1.80	0.04	0.02	0.94
	Highly correlated bands with Pan	3.99	5.19	0.08	0.05	0.87
	All bands	6.23	6.38	0.27	0.10	0.66
SFIM	Bands in Pan spectral range	1.83	1.27	0.04	0.05	0.91
	Highly correlated bands with Pan	2.38	2.72	0.06	0.05	0.89
	All bands	5.40	4.23	0.12	0.08	0.81
GLP-HPM	Bands in Pan spectral range	1.98	1.71	0.05	0.04	0.91
	Highly correlated bands with Pan	2.48	3.24	0.09	0.05	0.87
	All bands	5.50	4.58	0.18	0.08	0.76
CNMF	Bands in Pan spectral range	2.36	3.15	0.05	0.04	0.91
	Highly correlated bands with Pan	3.34	7.34	0.14	0.05	0.82
	All bands	10.39	15.70	0.36	0.18	0.53
HySure	Bands in Pan spectral range	2.27	2.96	0.19	0.17	0.68
	Highly correlated bands with Pan	4.80	3.29	0.17	0.16	0.70
	All bands	15.21	4.97	0.17	0.25	0.63

Table 2. 6 Measured quality indices for the Clovis dataset.

Fusion method	Selected bands for fusion	ERGAS	SAM	D_λ	D_S	QNR
PCA	Bands in Pan spectral range	2.23	1.87	0.05	0.03	0.92
	Highly correlated bands with Pan	2.77	3.62	0.10	0.10	0.81
	All bands	8.47	10.56	0.49	0.21	0.40
GS	Bands in Pan spectral range	2.24	1.73	0.03	0.04	0.93
	Highly correlated bands with Pan	1.81	5.01	0.13	0.11	0.78
	All bands	7.29	8.79	0.44	0.21	0.44
GSA	Bands in Pan spectral range	2.16	2.23	0.06	0.04	0.91
	Highly correlated bands with Pan	2.16	2.55	0.10	0.10	0.81
	All bands	2.86	3.24	0.26	0.17	0.61
UNB-Pansharp	Bands in Pan spectral range	2.16	1.08	0.06	0.06	0.88
	Highly correlated bands with Pan	2.34	1.28	0.09	0.06	0.85
	All bands	2.92	2.48	0.30	0.17	0.58
SFIM	Bands in Pan spectral range	1.60	1.02	0.05	0.09	0.86
	Highly correlated bands with Pan	1.82	1.37	0.11	0.11	0.79
	All bands	2.84	2.78	0.24	0.18	0.62
GLP-HPM	Bands in Pan spectral range	1.78	1.38	0.04	0.08	0.89
	Highly correlated bands with Pan	1.81	1.78	0.12	0.10	0.80
	All bands	2.61	3.20	0.26	0.18	0.60
CNMF	Bands in Pan spectral range	3.92	3.72	0.05	0.07	0.88
	Highly correlated bands with Pan	4.03	7.13	0.14	0.11	0.76
	All bands	6.81	13.36	0.29	0.18	0.58
HySure	Bands in Pan spectral range	1.94	2.73	0.25	0.16	0.63
	Highly correlated bands with Pan	2.13	4.08	0.19	0.20	0.65
	All bands	6.68	5.95	0.23	0.24	0.59

Another point of the investigation is to compare the flexibility of different methods in accepting the spectral bands from different regions of the spectrum beyond the Pan spectral range. Comparing the values of SAM, ERGAS, and D_λ for the three band selection scenarios in each method, shows that the spectral quality of the results for the PCA, GS, and CNMF methods varies greatly depending on the HS bands used, especially when all bands are used, the spectral distortion decreases dramatically. On the other hand, the spectral quality in GSA, UNB-Pansharp, SFIM, GLP-HPM, and HySure methods are less sensitive to the bands fed to them. The reason is that for GSA, UNB-Pansharp, and HySure methods, the spectral relationship between Pan and HS bands is estimated from the observed data in an optimum way to best fit the observations. More specifically, both the GSA and UNB-Pansharp methods are based on least square error, and as such the adaptive weights are assigned to different HS bands to obtain the low resolution version of the Pan band.

Another comparison can be conducted considering the spatial distortion parameter (D_S) for the three band selection scenarios in each method. Compared to the case when all band are used, using the bands highly correlated with Pan yields to less spatial distortion. This is mainly because the HS bands, especially those which are beyond the Pan spectral range with minimum correlation with the Pan band, contain some strip noises which affect the spatial quality of the fused images obtained from all bands. In addition, SFIM, GLP-HPM, and HySure methods yield to poor spatial qualities, as was expected from visual analysis.

To make the achievements of this study more applicable for real-world applications, we also provide the processing times consumed by each pan-sharpening method in Table 2.7. These values are from an Intel Core i5 3230 k machine with a clock speed of 2.6 GHz with 32 GB RAM. As can be seen, the methods designed specifically for HS pan-sharpening, CNMF and HySure, are expensive in terms of processing time, while the methods adapted from MS pan-sharpening are less time-consuming. In addition, by incorporating more bands in the fusion processing, the processing time increases, especially when all bands are used this increase is dramatically high for most of the methods.

Table 2. 7 Processing times of different pan-sharpening methods in different datasets (in seconds).

Dataset	Selected bands for fusion	PCA	GS	GSA	UNB-Pansharp	SFIM	GLPHPM	CNMF	HySure
Calgary	Bands in Pan spectral range	5.02	10.75	19.55	37.78	20.96	17.03	82.98	5259.44
	Highly correlated bands with Pan	21.79	94.54	221.05	81.13	63.47	93.76	299.47	5327.38
	All bands	2405.19	1331.10	1627.46	153.63	239.94	1044.74	449.47	5591.43
Tehran	Bands in Pan spectral range	5.84	11.79	22.04	43.69	15.44	18.57	90.75	5248.04
	Highly correlated bands with Pan	29.39	127.12	294.51	108.09	83.95	123.58	384.36	5434.97
	All bands	2638.61	1439.65	1984.50	158.62	1238.42	2341.02	1007.16	5887.97
Clovis	Bands in Pan spectral range	3.37	7.80	12.67	27.48	11.68	12.86	56.72	5174.86
	Highly correlated bands with Pan	15.43	65.51	165.40	60.20	48.82	68.70	223.53	5248.04
	All bands	1504.10	1025.62	1033.03	104.80	93.05	135.59	629.48	5291.62

Considering different band selection scenarios along with the processing time, we can conclude the following notes about different pan-sharpening methods:

• GS and PCA methods: these methods are the fastest methods and are sensitive to the bands used for fusion, especially using all bands for fusion, not only injects high spectral distortion, it increases the processing time dramatically. However, with selecting bands highly correlated with Pan, we could fuse more spectral bands beyond the Pan spectral range, without injecting high spectral distortion to the fused bands (based on Tables 4 to 6) with relatively low processing time (based on Tables 7).

• GSA and UNB-Pansharp methods: these methods are less sensitive to the bands used for fusion. Considering the measured quality indices in Tables 4 to 6 and the processing time reported in Table 2.7, selecting bands highly correlated with Pan, results in acceptable spectral and spatial qualities with acceptable processing load.

• SFIM and GLP-HPM methods: for these methods, the spatial quality is more affected by the band selection rather than the spectral quality. Selecting bands highly correlated with Pan yields to the results with less noise or in other words better spatial quality. However, compared to other methods, the spatial quality of the results of these methods is not satisfying for the pan-sharpening purposes.

• CNMF and HySure methods: the application of these methods is strongly limited by their heavy processing load. The spectral quality of the results of the CNMF method is highly sensitive to the bands used for it, especially using all bands for fusion, distorts the spectral quality highly. However, with selecting bands highly correlated with Pan, the number of fused bands increases without injecting high spectral distortion. About the HySure method, the spatial quality is not as good as the CS and UNB-Pansharp methods, while its processing time is significantly high regardless of the number of bands used for fusion. However, selecting bands highly correlated with Pan results in fused bands with superior spatial quality compared to when all bands are used.

2.3.3.3 Spectral Signatures

In this section, we provide further observation to see how the proposed adaptive band selection method affect the spectral signatures of different classes that exist in the original HS image. For this purpose, we plot the spectral signatures of different land cover types, vegetation, water, soil, and urban, extracted from the original HS image and the fused images obtained from bands highly correlated with Pan in Figures 2.4 to 2.6. To better compare the signatures, only the common bands were selected for visualization in Figures 2.4 to 2.6. The fused result which reconstructs the most similar and the closest signatures to the original ones can be considered as the result with the minimum spectral distortion. It is noteworthy that these signatures are obtained by manually selecting different areas containing different land cover types throughout the images, and by averaging the spectra obtained over those selected areas.

Comparing the signatures obtained from fused images with the original ones obtained from HS images in Figures 2.4 to 2.6 shows that generally, the signatures follow the trend of the original signatures quite well in different areas of the spectrum, regardless of the land cover type. In fact, although the digital values have changed, especially in PCA, GS and CNMF methods, the trend of the original signatures is still preserved; hence, there would be no problem in the identification procedure of these class types in applications such as classification. In summary, with adaptive band selection, we could incorporate more spectral bands in the pan-sharpening procedure, without distorting the signature in the visible range nor in other areas of the spectrum.

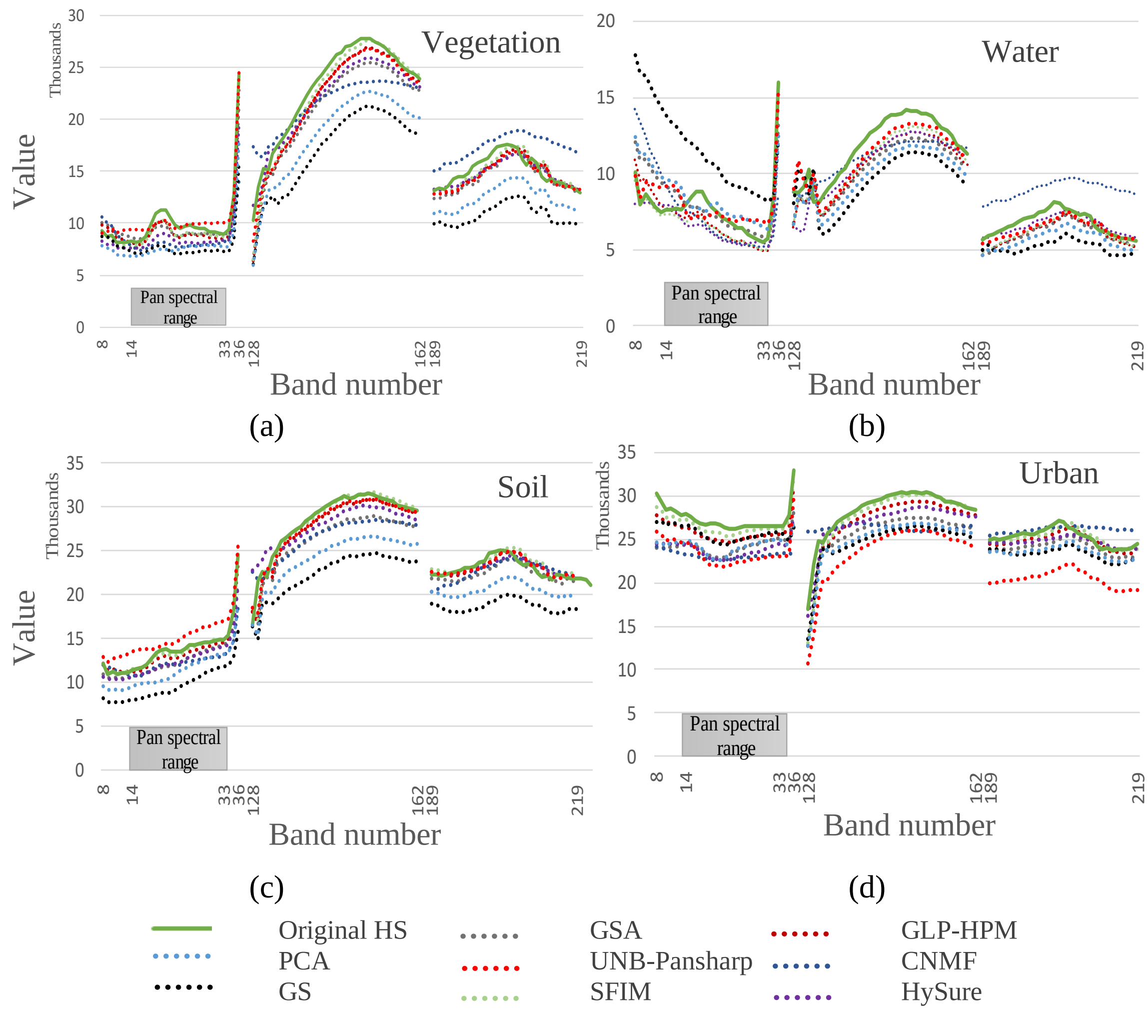


Figure 2. 4 Spectral signatures of a) vegetation, b) water, c) soil, and d) urban areas in Calgary dataset, extracted from the original HS and the fused images obtained using bands highly correlated with Pan in different methods.

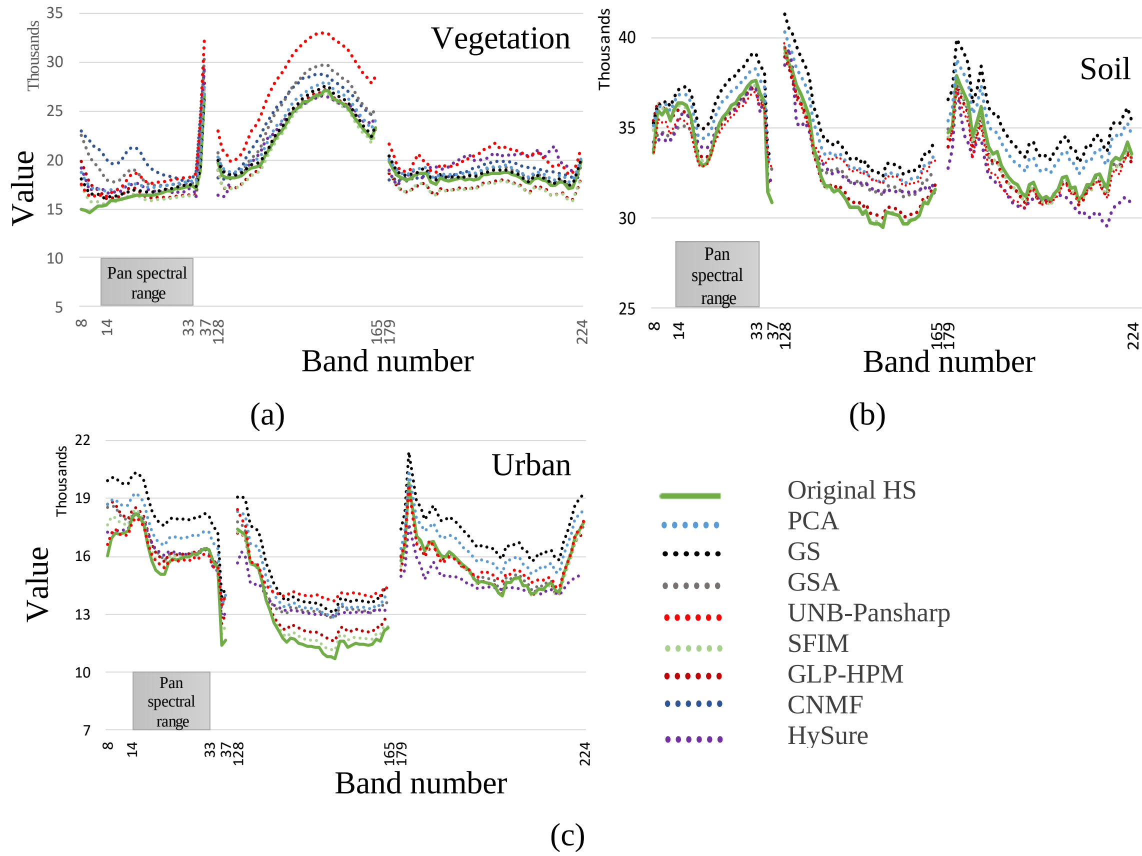


Figure 2. 5 Spectral signatures of a) vegetation, b) soil, and c) urban areas in Tehran dataset, extracted from the original HS and the fused images obtained using bands highly correlated with Pan in different methods.

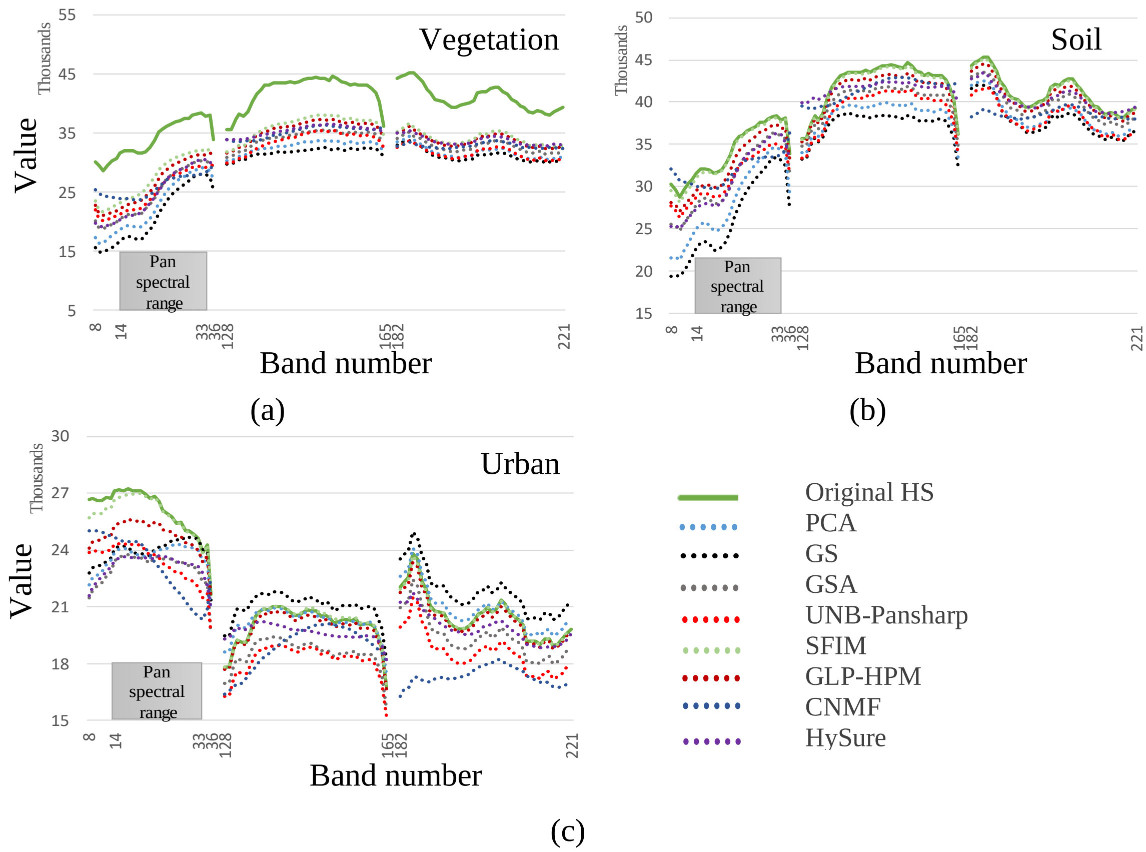


Figure 2. 6 Spectral signatures of a) vegetation, b) soil, and c) urban areas in Clovis dataset, extracted from the original HS and the fused images obtained using bands highly correlated with Pan in different methods.

2.4 Conclusions

The HS pan-sharpening so far was limited to only the bands that share a common spectral range with the Pan band, which results in a limited number of fused bands compared to the total number of HS bands. In this study, an adaptive band selection method for HS pan-sharpening is proposed, which involves selecting those HS bands that are highly correlated with the Pan band throughout the whole spectral range, instead of selecting the bands that fall within the Pan spectral range. To test the proposed approach, the HS and Pan bands of three different datasets acquired by EO-1 satellite were fused using PCA, GS, GSA, SFIM, GLP-HPM, and UNB-Pansharp methods, adapted from MS pan-sharpening, and CNMF and HySure, as two methods designed specifically for HS pan-sharpening. For each dataset and each method, three fused images were generated using: 1) the HS bands within the Pan spectral range, 2) the HS bands that are highly correlated with Pan, and 3) all the HS bands.

Results show that with the proposed adaptive band selection, it is possible to incorporate more HS bands (more than those within the Pan spectral range) in the pan-sharpening procedure without adding high spectral distortions to the results. In addition, the trend of the spectral signatures obtained from the adaptive band selection was very close to those obtained from the original HS, which guarantees that the proposed approach would not affect the identification procedure of different land cover types in applications such as classification.

Moreover, it was observed that among the investigated methods, PCA and GS methods are highly sensitive to the bands that are selected for the fusion, especially using all HS bands results in high spectral distortions. Thus, it is crucially suggested to use the proposed adaptive band selection instead of using all bands for these methods. The GSA and UNB-Pansharp methods, with acceptable processing time, are flexible to accept the bands from different parts of the spectral range, mainly due to the functionality of least square error in their modelling.

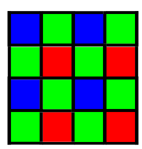
For SFIM, and GLP-HPM, generally the spatial quality is not as high as CS methods, and the spectral quality is not affected that much by different band selection. However, the spatial quality of the results obtained from bands highly correlated with Pan is superior to the results obtained from all bands. Also, for CNMF and HySure methods, adaptive band selection provides a higher number of fused bands compared to when the bands within Pan spectral range are used, while higher spectral and spatial qualities are obtained compared to using all bands. However, the application of these two methods is strongly limited by their heavy processing load. The effect of the proposed adaptive band selection can be further investigated in real-world applications such as classification which is considered as the future work of the current study.

3 RGBW Demosaicking Using a Collaborative Interpolation Between Panchromatic and Colour Pixels

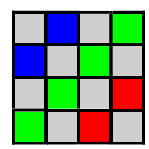
3.1 Introduction

Single-chip colour cameras are widely used in the imaging industry due to their lower manufacturing costs compared to three-chip colour cameras. A single-chip camera is equipped with a Colour Filter Array (CFA), containing red (R), green (G) and blue (B) filter elements, placed on top of the sensor. Each filter element receives only one out of the three primary lights of R, G or B, while the other two are blocked. This results in a single-channel image known as a raw image. To reconstruct a full resolution three-channel colour image, a procedure known as demosaicking (or debayering) is employed to the raw image that involves interpolating the missing pixel values using neighbouring values. Based on the arrangement of the R, G and B filter elements, different CFAs are available (Losson et al., 2010). Bayer CFA (Figure 3.1 (a)) is the most commonly used one in the industry (Bayer, 1976) for which various demosaicking methods have been developed and studied (a review of them can be found in (Losson et al., 2010; Menon & Calvagno, 2011; Monno et al., 2017)).

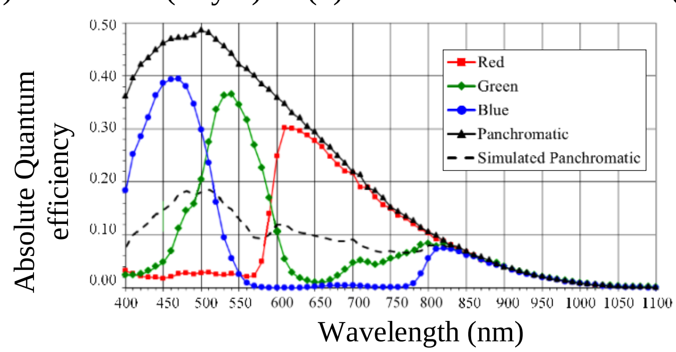
The problem with RGB CFAs is that the received light is significantly reduced due to the filtering at each filter element, resulting in a lower Signal to Noise Ratio (SNR). This problem is more severe in low lighting conditions or with shorter exposure times. To improve the light sensitivity, Kodak embedded some transparent filter elements, known as panchromatic or white (W), in the RGB CFA and produced an RGBW CFA. Kodak then built a sensor equipped with such a CFA, called Truesense sparse CFA, which is known as CFA2.0 in the literature (Figure 3.1 (b)). It generates a raw image where half of the pixels are panchromatic and the other half belongs to R, G, and B. Since panchromatic pixels are generated with no blockage of the visible spectrum energy, they provide higher sensitivity than RGB colour pixels which finally leads to improvement of the image SNR (H. Kim et al., 2020). In a comparative study (Kwan & Larkin, 2019) it is shown that having more panchromatic pixels improves the image quality compared to the RGB CFAs, especially in low lighting conditions. Hereafter, we use ‘Truesense’ to refer to the sensor equipped with the Truesense sparse CFA, while ‘CFA2.0’ mainly refers to the raw images obtained from such sensors, and ‘Bayer image’ refers to the raw image obtained from a Bayer CFA.



(a) RGB CFA (Bayer)



(b) RGBW CFA or CFA2.0 (Truesense)



(c) Quantum efficiency of Kodak Truesense sensor (plot adopted from (AND9180/D-TRUESENSE Sparse Color Filter Pattern, 2014)). The plot for the simulated panchromatic is obtained by averaging Red, Green and Blue plots.

Figure 3. 1 Comparing RGB and RGBW CFAs and the related spectral responses.

The higher sensitivity of RGBW CFA compared to the RGB CFA can be better explained by comparing the spectral response of the related sensors through Quantum Efficiency (QE) depicted in Figure 3.1 (c). QE is the ratio of photogenerated electrons captured by a pixel to the number of incident photons over a period of time (AND9180/D-TRUESENSE Sparse Color Filter Pattern, 2014). As can be seen in Figure 3.1 (c), the spectral range of the panchromatic pixels is wider than that of any colour pixels, and the spectral response of the panchromatic pixels in total is more than twice that of the individual colour pixels.

The RGBW CFA is designed in a way that one can extract the standard Bayer image simply by downsampling colour pixels to a half resolution level. This extracted Bayer image can then be demosaicked using any available method for the Bayer demosaicking. Accordingly, Kodak (Hamilton & Compton, 2007) suggests interpolating colour and panchromatic images in different resolution levels.

At full resolution level, the panchromatic image is interpolated using panchromatic pixels without involving colour pixels. This generates a High Resolution Panchromatic (HRP) image. At half resolution level, the extracted Bayer image is demosaicked. The panchromatic pixels are used directly for G interpolation and the obtained G channel is then used for R and B interpolations. This generates a Low Resolution Colour (LRC) image. The full resolution colour image is finally obtained by integrating the HRP and LRC images. This integration can be viewed as a simple fusion approach, where the differences between LRC and the downsampled HRP are calculated, then upsampled, and finally added to the HRP image. Later, in a white paper, Bayer CFA and RGBW CFA were compared using the images taken by a Truesense sensor (AND9180/D-

TRUESENSE Sparse Color Filter Pattern, 2014). It was demonstrated that the light sensitivity was significantly increased (between 2 and 4 times) compared to the Bayer CFA while maintaining overall colour fidelity. However, our observations show that the resultant images suffer from visible colour leakages and zipper artifacts especially in the areas related to high frequency information, i.e. the spatial resolution is distorted.

Kwan et al. (2017) proposed to apply image fusion techniques in the procedure of CFA2.0 demosaicking. They followed a similar procedure as in Kodak's method (Hamilton & Compton, 2007) but used other demosaicking methods to generate HRP and LRC images and other widely used fusion techniques to fuse them. We refer to this approach as '*II+fusion*', where *II* stands for independent interpolation. By independent interpolation, we mean that HRP is generated independently of the colour pixels and the LRC image is demosaicked without involving the panchromatic pixels. They concluded that their preliminary results do not show significant improvement compared to Kodak's method mainly because they used a very simple algorithm for panchromatic interpolation. However, we think that the reason is related not only to panchromatic interpolation but also to demosaicking the colour image at half resolution level. Among the Bayer demosaicking methods, it is shown in (Monno et al., 2017) that the Adaptive Residual Interpolation (ARI) outperforms the state-of-the-art methods. ARI uses the guided filter technique (He et al., 2013) to guide interpolating the G missing pixels using R and B available pixels and to guide R and B interpolations using the G channel.

Besides, other methods perform demosaicking in one resolution level. Their approach mainly focuses on estimating the luminance and chrominance components either in the frequency domain (Rafinazari, 2017; J. Wang et al., 2011), or in the spatial domain through the so-called universal methods (Condat, 2009, 2012; Singh & Singh, 2012; C. Zhang et al., 2016).

Reviewing the literature, it can be seen that some methods involve two main procedures, i.e. demosaicking and fusion (Hamilton & Compton, 2007; Kwan et al., 2017; Kwan & Chou, 2019), while others involve only demosaicking (Condat, 2009, 2012; Rafinazari, 2017; Singh & Singh, 2012; J. Wang et al., 2011; C. Zhang et al., 2016). This study aims at improving the performance of the demosaicking-and-fusion method by proposing a Collaborative Interpolation (CI) between panchromatic and colour pixels followed by a fusion method. We refer to the whole procedure as '*CI+fusion*', however, our main contribution concentrates on CI. Using CI, we treat RGB pixels as an auxiliary observation that can be incorporated in the panchromatic interpolation and, vice versa; we treat the panchromatic pixels as an auxiliary observation that can be incorporated in interpolating R, G, and B channels. In summary,

- for generating the HRP image, instead of a common interpolation that uses only observed panchromatic pixels, we propose to incorporate RGB values to guide panchromatic interpolation.
- for generating the LRC image, we propose to incorporate the panchromatic pixels to guide interpolating all the colour channels.

With such modifications, first, we generate a panchromatic image with the highest possible spatial resolution (i.e. fine spatial details). We then inject such spatial details into the colour image not only through fusion (at full resolution level) but also through demosaicking the colour image (at half resolution level).

The procedure we used for CI is inspired by the ARI method (Monno et al., 2017). The idea of the guided filter in ARI is well aligned with our purpose of using each of the colour/panchromatic pixels to guide the interpolation of the other one. However, we cannot directly use ARI, since it is designed for the Bayer CFAs; thus, we need to adapt it for RGBW CFA and to modify it in a way so that we can incorporate the panchromatic pixels in the R, G, and B interpolations. More specifically, we have adapted ARI for interpolating panchromatic and G channels, and for R and B channels, we have developed an iterative version of Minimized-Laplacian Residual Interpolation (MLRI). The details of such modifications will be provided in Section 3.2.

For the evaluation, we have designed two scenarios: 1) we followed the same routine in the literature by simulating CFA2.0 images using Kodak benchmark dataset, 2) we captured real CFA2.0 images by a camera equipped with a Truesense sensor, which besides raw images provides its own demosaicked images. The latter scenario is missing in the experiments found in the literature, and is more favourable as simulations cannot fully reflect the real-world imaging conditions. In both scenarios, we compare our method to the modified version of Kwan et al. (2017), named II+fusion, and to Condat's method with two regularizations, Tikhonov (Tik; Condat, 2009), and total variation (TV) minimization (Condat, 2014, 2012). The comparison is drawn using the colour peak signal-to-noise ratio (CPSNR) and CIE-LAB errors. For the second scenario, in addition to the mentioned methods, we have compared our results with the camera's results using SNR, CIE-LAB error, and the proposed 'Channel Discrepancy'. The visual comparison of the results shows that our approach produces the least colour and zipper artifacts along edges, the least colour noise and better preserves the fine details. The quantitative evaluations further confirm that our proposed method outperforms the investigated methods as well as the method employed by the commercial camera.

The rest of this chapter is organized as follows. Section 3.2 first gives a detailed description of the proposed method, including the idea of CI as well as three main phases of the procedure. The results are presented and discussed in Section 3.3. Finally, Section 3.4 concludes the chapter.

3.2 Proposed CFA2.0 Demosaicking Procedure

The general processing pipeline of the proposed method is illustrated in Figure 3.2, which comprises three main phases: Phase (i) is to generate the HRP image using the mosaiced panchromatic image (P) and the sparse mosaicked colour image (C). The details of this phase are provided in Section 3.2.2. Phase (ii) involves generating the LRC image by demosaicking the standard Bayer image (C') at half resolution, in which the panchromatic pixels from the Low Resolution Panchromatic (LRP) image are incorporated. The details of this phase are provided in Section 3.2.3. To obtain C' and LRP images, the images P and C are downsampled by averaging two pixels of the same colour in a 2x2 neighbourhood. Phase (iii) consists of fusing the HRP and LRC images generated in the previous phases to reconstruct the full resolution colour image (the details of this phase are provided in Section 3.2.4). In the remainder of this section, first, the main idea of CI is explained, and then, a detailed description of the three phases is provided.

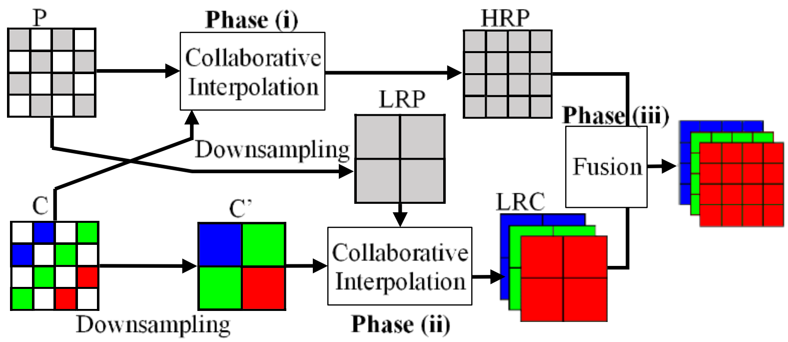


Figure 3. 2 The proposed procedure for CFA2.0 demosaicking.

3.2.1 Main Idea of CI

The main idea of the proposed CI is to use colour pixels to guide interpolation of the panchromatic image at the high resolution level and to use panchromatic pixels to guide demosaicking of the colour image at half resolution level. The idea of CI comes from the fact that for a target scene, two different observations from two different sources, panchromatic and colour sensors, are available. Based on (Kimmel, 1999), a simplified model for an image formation can be considered as the projection of the normal vector ($\vec{N}$) of an object's surface onto the incident light vector ($\vec{l}$) multiplied by the reflection coefficient ρ. Considering that ρ is related to the object material and hence is different for different colour channels, the intensity I received by a pixel at the position (i,j) can be written as (Kimmel, 1999):

$$I^k(i,j) = \rho^k(i,j).\langle\vec{N}(i,j),\vec{l}\rangle, \quad k \in \{R,G,B\}. \tag{3.1}$$

Therefore, the differences between any two colour channels, k and k' at the pixel (i,j) can be expressed as (Losson et al., 2010):

$$D^{k,k'}(i,j) = I^k(i,j) - I^{k'}(i,j)$$

$$= \rho^k(i,j).\langle\vec{N}(i,j),\vec{l}\rangle - \rho^{k'}(i,j).\langle\vec{N}(i,j),\vec{l}\rangle$$

$$= const. \tag{3.2}$$

Since the material is constant within an object, the reflection coefficients, ρ^k, are assumed to be constant within a neighbourhood ($\omega_{i,j}$) of the pixel (i, j). Moreover, the normal vector and the light vector are assumed to be locally constant. Therefore, the differences between any two channels in the neighbourhood $\omega_{i,j}$ can be considered constant (Losson et al., 2010):

$$\rho^k_{\omega_{i,j}}.\langle \vec{N}_{\omega_{i,j}}, \vec{l} \rangle - \rho^{k\prime}_{\omega_{i,j}}.\langle \vec{N}_{\omega_{i,j}}, \vec{l} \rangle = const. \tag{3.3}$$

The colour difference constancy is the main assumption in the most widely used RGB demosaicking procedures (Losson et al., 2010). However, we can consider equation (3.3) to be valid between panchromatic and any colour channels in RGBW CFAs (i.e. $k, k' =\in \{R, G, B, P\}^2$).

Although the target scene is sampled differently by each panchromatic and colour sensors, equation (3.3) shows that the colour and panchromatic images are highly correlated, at least in local windows. Spatial and spectral correlations, based on (Losson et al., 2010), are two main rules that should be enforced to improve the demosaicking results. However, such correlations between panchromatic and colour pixels have been neglected in the procedures proposed by Kodak (Hamilton & Compton, 2007) and (Kwan et al., 2017).

Based on equation (3.3), we can establish a model that relates the panchromatic and colour images in local windows. If we write equation (3.3) for $k \in \{P\}$ and $k' \in \{R, G, B\}$, we will have three equations that by adding them together, we get:

$$\rho^P_{\omega_{i,j}}.\langle \vec{N}_{\omega_{i,j}}, \vec{l} \rangle - \frac{1}{3}\sum_k \rho^k_{\omega_{i,j}}.\langle \vec{N}_{\omega_{i,j}}, \vec{l} \rangle = const. \tag{3.4}$$

Note that in a CFA2.0 image, $\sum_k \rho^k_{\omega_{i,j}} . \langle \vec{N}_{\omega_{i,j}}, \vec{l} \rangle$ or simply $\sum_k I^k_{\omega_{i,j}}$, is equivalent to the sparse mosaicked colour image C in Figure 3.2. Based on equation (3.4), we can assume that the panchromatic image, at high resolution level (shown by the superscript H), is a linear transformation of the image C in a window ω, i.e.

$$P^H(i,j) = a_\omega . C^H(i,j) + b_\omega, \ \forall (i,j) \epsilon \omega. \tag{3.5}$$

The coefficients a_ω and b_ω are unknown and assumed to be constant for all the pixels in ω. Similarly, equation (3.3) can be written at a low resolution level (shown by the superscript L) with $k \in \{R, G, B\}$ and $k' \in \{P\}$. Thus, we can assume that each low resolution colour channel is a linear transformation of the low resolution panchromatic image in ω with the unknown coefficients of a'_ω and b'_ω:

$$C^{L,k}(i,j) = a'_\omega . P^L(i,j) + b'_\omega, \forall k \epsilon \{R, G, B\} \ and \ \forall (i,j) \epsilon \omega. \tag{3.6}$$

Here, we only provide the solution for equation (3.5). Similarly, the unknowns of equation (3.6) are achievable. To determine the unknowns, based on the guided filtering (GF) technique (He et al., 2013), the following cost function should be minimized:

$$E(a_\omega, b_\omega) = \sum_{i,j \in \omega} (P^H(i,j) - a_\omega . C^H(i,j) - b_\omega)^2. \tag{3.7}$$

Equation (3.7) enforces that the difference between the panchromatic image and the colour image is the lowest possible amount in the window ω, whereas equation (3.5) and (3.6) are maintained. The unknown coefficients are then given by:

$$a_\omega = (\frac{1}{|\omega|} \sum_{i,j \in \omega} C^H(i,j) . P^H(i,j) - \mu_{C_\omega} . \mu_{P_\omega}) / \sigma^2_{P_\omega} \tag{3.8}$$

$$b_\omega = \mu_{P_\omega} - a_\omega . \mu_{C_\omega}, \tag{3.9}$$

where μ_{C_ω} and μ_{P_ω} are the mean values of the colour and the panchromatic images in ω, respectively, and $\sigma^2_{P_\omega}$ is the variance of the panchromatic image in ω, and $|\omega|$ is the total number of pixels in ω. Since the unknowns are calculated in sliding windows, for one specific pixel at (p, q), as far as it exists in the sliding window, several coefficients can be achieved. The final coefficients $(\bar{a}, \bar{b})$ are then calculated as proposed in (Kiku et al., 2016) through a weighted average of all the obtained coefficients:

$$\bar{a}(p,q) = \frac{\sum_{t=1}^{t=n} W_{\omega_t}.a_{\omega_t}}{\sum_{t=1}^{t=n} W_{\omega_t}}, \bar{b}(p,q) = \frac{\sum_{t=1}^{t=n} W_{\omega_t}.b_{\omega_t}}{\sum_{t=1}^{t=n} W_{\omega_t}}$$

$$W_{\omega} = 1/\frac{1}{|\omega|}\sum_{x,y\in\omega}(P^H(i,j) - a_{\omega}.C^H(i,j) - b_{\omega})^2, \tag{3.10}$$

where n is the total number of all the windows that contain the specific pixel at (p, q), and W is the weight which is calculated based on the residual cost; thus, the coefficients that result in high residuals contribute less to the calculations of $\bar{a}$ and $\bar{b}$. Once the linear coefficients are calculated for equations (3.5) and (3.6), we calculate an estimate of the panchromatic pixel $(\hat{P}^H)$ and an estimate of the colour pixel $(\hat{C}^L)$ through:

$$\hat{P}^H(i,j) = \bar{a}\,(i,j).C^H(i,j) + \bar{b}\,(i,j), \tag{3.11}$$

$$\hat{C}^{L,k}(i,j) = \bar{a}'\,(i,j).P^L(i,j) + \bar{b}'\,(i,j), \quad \forall k\epsilon\{R,G,B\}. \tag{3.12}$$

Equations (3.11) and (3.12) are the main core of the collaborative interpolation process. Like residual-interpolation-based methods in RGB demosaicking (Kiku et al., 2016; Monno et al., 2017), here, we can find the residuals for the missing panchromatic pixels. To do so, the residuals between the real and estimated panchromatic pixels should be calculated. Equation (3.11) gives the estimated panchromatic pixels using the colour pixels, C^H (more details are given in Section 3.3.2). Similarly, the residuals between the real and estimated colour pixels, can be used to find the residuals for the missing colour pixels. Equation (3.12) gives the estimated colour pixels using the panchromatic pixels (more details are given in Section 3.2.3). The following equation proves that the colour difference constancy in equation (3.3) is valid for the residual calculation as well considering that $\bar{a}_\omega$ and $\bar{b}_\omega$ are constant in ω:

$$\rho_\omega^k.\langle \vec{N}_\omega, \vec{l} \rangle - \bar{a}_\omega.\rho_\omega^{k\prime}.\langle \vec{N}_\omega, \vec{l} \rangle - \bar{b}_\omega = const. \tag{3.13}$$

3.2.2 Phase (i): Generating HRP Image Using CI

For the panchromatic interpolation, we adapt the G interpolation in ARI that consists of three steps: (i) iterative directional interpolation by two methods, Residual Interpolation (RI; Kiku et al., 2013), and Minimized-Laplacian RI (MLRI; Kiku et al., 2016); (ii) adaptive selection of iteration number; and (iii) adaptive combining. For each RI/MLRI method, iterative directional interpolations are performed. This generates k images for each method, where k is the maximum iteration number. At each pixel, the best value is selected using a criterion. Finally, four interpolated images are obtained and combined to generate the HRP image (Monno et al., 2017).

To adapt ARI for the panchromatic interpolation, the question was which colour channel and in what order should be utilized as the guidance to produce the highest possible quality. This was answered by equations (3.4) and (3.5) which show that the linear model established between panchromatic and colour pixels takes all the colour pixels at once through image C. This was also confirmed through a collaborative project with Canada's National Optics Institute (INO). It was observed that the best quality is achieved when all colour channels (R, G and B) are considered as a single channel. This can be further explained by the spatial resolution of the involved images. In Phase (i), our main concern is to retain the highest possible spatial resolution and we ignore reconstructing the precise colour (Phase (ii) and (iii) are involved with retrieving and preserving the colour). Taking any R, G, or B channels separately as the guidance would result in degrading the spatial resolution of the HRP image since each colour channel has a lower spatial resolution than that of the panchromatic image. However, the spatial resolution of the image C is equal to the spatial resolution of the panchromatic image. Therefore, we propose to apply inter-channel residual interpolation instead of the common residual interpolation in the first step of ARI. Figure 3.3 shows the overall steps of CI applied to the panchromatic image using the image C. Although the main modification is related to step (i), we provide the details of all the steps in the following sections.

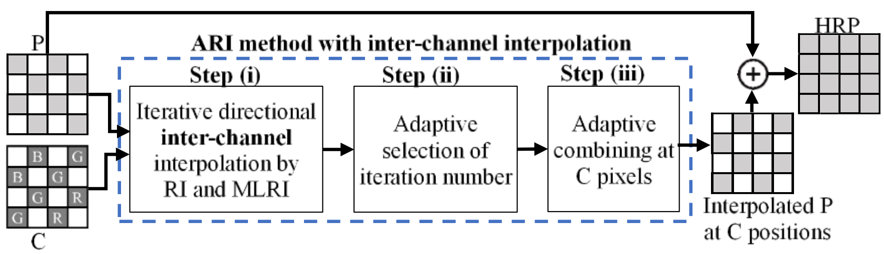


Figure 3. 3 The overall flow of CI applied to the panchromatic channel (P).

Step (i): Iterative Directional Inter-Channel Interpolation by RI and MLRI

The residual in ARI is defined as the difference between an estimated and the related observed pixel values. The iterative directional interpolation of residuals starts with initializing the missing pixels of C and P channels by horizontal/vertical interpolations, which involves averaging two neighbouring available values in the rows/columns. Based on Figure 3.3, the neighbouring available values in image C belong to different colour channels. This shows why we describe such interpolation as inter-channel interpolation. We show the initial interpolation results by $\tilde{P}_0$ and $\tilde{C}_0$ for the panchromatic and colour images, respectively. At each iteration k, the estimated values ($\hat{P}$ and $\hat{C}$) are calculated in a local window around a pixel (i, j) by linear transformations of the interpolated images at the previous iteration ($\tilde{P}_{k-1}$ and $\tilde{C}_{k-1}$):

$$\hat{P}_k^D(i,j) = a_k^P(i,j).\tilde{C}_{k-1}(i,j) + b_k^P(i,j),$$

$$\hat{C}_k^D(i,j) = a_k^C(i,j).\tilde{P}_{k-1}(i,j) + b_k^C(i,j),$$

$$D \in \{h\text{: } horizontal, v\text{: } vertical\}, \tag{3.14}$$

where (a^P, b^P) and (a^C, b^C) are the unknown coefficients at each local window. For the first iteration, $k = 1$, the initialized images ($\tilde{P}_0$ and $\tilde{C}_0$) are used as guidance. To calculate the unknowns, equations (3.8) to (3.10) are used, which are applicable for both horizontal and vertical directions. Now, to obtain the residuals, the estimated pixel values are subtracted from the ones which were originally available in the input images:

$$r_k^{P^D}(i,j) = P(i,j) - \hat{P}_k^D\,(i,j), (i,j) \in \chi^P,$$

$$r_k^{C^D}(i,j) = C(i,j) - \hat{C}_k^D(i,j), (i,j) \in \chi^C, \tag{3.15}$$

in which, χ^P and χ^C are the sets of positions where the panchromatic and colour pixels are originally available. The missing residuals are calculated by averaging the two neighbouring residuals in the horizontal and vertical directions,

$$r_k^{P^h}(i,j) = (r_k^{P^h}(i-1,j) + r_k^{P^h}\,(i+1,j))/2, (i,j) \in \chi^C$$

$$r_k^{C^h}(i,j) = (r_k^{C^h}(i-1,j) + r_k^{C^h}(i+1,j))/2, (i,j) \in \chi^P \tag{3.16}$$

$$r_k^{P^v}(i,j) = (r_k^{P^v}(i,j-1) + r_k^{P^v}\,(i,j+1))/2, (i,j) \in \chi^C$$

$$r_k^{C^v}(i,j) = (r_k^{C^v}(i,j-1) + r_k^{C^v}\,(i,j+1))/2, (i,j) \in \chi^P \tag{3.17}$$

The interpolated residuals are then added to the estimated pixels to generate the final interpolated image for each iteration:

$$\tilde{P}_k^D(i,j) = r_k^{P^D}(i,j) + \hat{P}_k^D(i,j), \quad (i,j) \in \chi^C$$

$$\tilde{C}_k^D(i,j) = r_k^{C^D}(i,j) + \hat{C}_k^D(i,j), \quad (i,j) \in \chi^P. \tag{3.18}$$

These interpolated images are used in equation (3.14) as the new guidance set, $\tilde{P}$ and $\tilde{C}$, for which a new set of coefficients are calculated through GF. The window size in GF is enlarged at each iteration. We set the window size for RI initially as 1×2 (height × width) and which is enlarged by 1 for the height and by 2 for the width at each iteration. This continues until the maximum iteration number, $k = 11$ as in (Monno et al., 2017), is reached.

For MLRI, the procedure is quite similar. The main difference is that the coefficients a^P and a^C are calculated differently, i.e. by locally minimizing the directional Laplacian of the residuals instead of the residuals. Thus, the coefficients for each local window (ω) are calculated by minimizing the following modified cost functions at each iteration k (Kiku et al., 2014):

$$a_{\omega,k}^{P,D} = arg\,min \sum_{i,j\in\omega} \left(\widetilde{\nabla}^2 P_k^D\,(i,j) - \widetilde{\nabla}^2 \hat{P}_k^D\,(i,j)\right)^2$$

$$= arg\,min \sum_{i,j\in\omega} \left(\widetilde{\nabla}^2 P_k^D\,(i,j) - a_{\omega,k}^{P,D}\,.\widetilde{\nabla}^2 \tilde{C}_{k-1}^D\,(i,j)\right)^2, \tag{3.19}$$

where $\widetilde{\nabla}^2(.)$ is the approximated Laplacian operator, with a sparse version of Laplacian kernel, $L = [-1\ \ 0\ \ 2\ \ 0\ -1]$, and its transposed version for the directional convolutions (Kiku et al., 2014). Similarly, the coefficient a^C can be calculated. The coefficients b^P and b^Ccan still be obtained using equation (3.9). The final value of the coefficients, obtained from different sliding windows, is calculated based on equation (3.10). The rest of the procedure is the same as the iterative RI method. In our experience, the window size (height × width) of GF for MLRI is initially set as 2×6 and is enlarged by 1 for the height and by2 for the width at each iteration.

Step (ii): Adaptive Selection of the Iteration Number

For this step, we follow the same routine as in (Monno et al., 2017). A criterion image is generated for each iteration based on the differences between the estimated and the interpolated pixel values at two consecutive steps (Monno et al., 2017):

$$d_k^P\ (i,j) = \hat{P}_k(i,j) - \tilde{P}_{k-1}(i,j)$$

$$d_k^C\ (i,j) = \hat{C}_k(i,j) - \tilde{C}_{k-1}(i,j). \tag{3.20}$$

Then, the criterion image at k-th iteration, crt_k, is calculated using (Monno et al., 2017):

$$crt_k(i,j) = \left(d_k(i,j)\right)^2 . \delta d_k(i,j),$$

$$d_k(i,j) = |d_k^P\ (i,j)| + \left|d_k^C\ (i,j)\right| ,$$

$$\delta d_k(i,j) = |d_k^P\ (i-1,j) - d_k^P\ (i+1,j)| + \left|d_k^C\ (i-1,j) - d_k^C\ (i+1,j)\right| \tag{3.21}$$

Then, a Gaussian smoothing kernel with $\sigma = 2$, g(.), is convolved with the criterion image of each iteration, resulting in smoothed criterion images, S_k, (Monno et al., 2017):

$$S_k(i,j) = \ g\left(crt_k(i,j)\right). \tag{3.22}$$

Therefore, if the maximum iteration number is set to 11, then 11 smoothed criterion images will be generated. Finally, for each pixel, the minimum value among the 11 smoothed criterion images is considered as the optimum value.

Step (iii): Adaptive Combining

For this step, we follow the same routine as in (Monno et al., 2017). After selecting the optimum value for each pixel, four interpolated images are available: two directional images from RI ($\tilde{P}_{RI}^{H}, \tilde{P}_{RI}^{V}$), and two directional images from MLRI ($\tilde{P}_{MLRI}^{H}, \tilde{P}_{MLRI}^{V}$). Hereafter, we ignore (i, j) to simplify the notation. The final interpolated HRP image is then calculated by combining the four obtained images through a weighted average (Monno et al., 2017):

$$\tilde{P} = \frac{(w_{RI}^{H}.\tilde{P}_{RI}^{H} + w_{RI}^{V}.\tilde{P}_{RI}^{V} + w_{MLRI}^{H}.\tilde{P}_{MLRI}^{H} + w_{RI}^{V}.\tilde{P}_{MLRI}^{V})}{(w_{RI}^{H} + w_{RI}^{V} + w_{MLRI}^{H} + w_{MLRI}^{V})}. \tag{3.23}$$

The weights are calculated based on the four smoothed criterion images (S_{RI}^{H}, S_{RI}^{V}, $S_{MLRI}^{H}, S_{MLRI}^{V}$) as:

$$w_{RI}^{H} = \frac{1}{S_{RI}^{H}},\ w_{RI}^{V} = \frac{1}{S_{RI}^{V}},\ w_{MLRI}^{H} = \frac{1}{S_{MLRI}^{H}},\ w_{MLRI}^{V} = \frac{1}{S_{MLRI}^{V}}. \tag{3.24}$$

3.2.3 Phase (ii): Generating LRC Image Using CI

Referring to Figure 3.2, the sparse mosaicked colour image C is downsampled to the Bayer image C’. Phase (ii) involves demosaicking this Bayer image, i.e. interpolation of low resolution R, G, and B channels. In CI, we also need to input the LRP image to involve the panchromatic pixels as well. For simplicity of notation, we refer to the channels of the LRC and LRP images again as R, G, B and P, respectively.

3.2.3.1 Interpolation of the G Channel

Interpolating the G channel can benefit from P pixels as well as R and B pixels. The outline of CI applied to the G channel is illustrated in Figure 3.4. For the Bayer image, we utilize the standard ARI method. In this case, the R and B pixels are used, and the G values are interpolated at the position of R and B pixels ($\tilde{G}^{RB}$). On the other hand, we can simply adapt the ARI method to utilize P pixels, instead of R and B pixels, which gives $\tilde{G}^{P}$. Finally, we combine $\tilde{G}^{RB}$ and $\tilde{G}^{P}$ using a weighted averaging.

The G interpolation in ARI consists of the same three steps described in Sections 3.2.2.1 to 3.2.2.3: iterative directional interpolation by RI and MLRI, adaptive selection of iteration number, and adaptive combining at R/B pixels. The window size in GF and the maximum iteration number is set the same as described for panchromatic interpolation. The G values are interpolated at R positions using the R pixels and similarly at B positions using the B pixels. Finally, the interpolated G pixels at different positions are put together to generate $\tilde{G}^{RB}$.

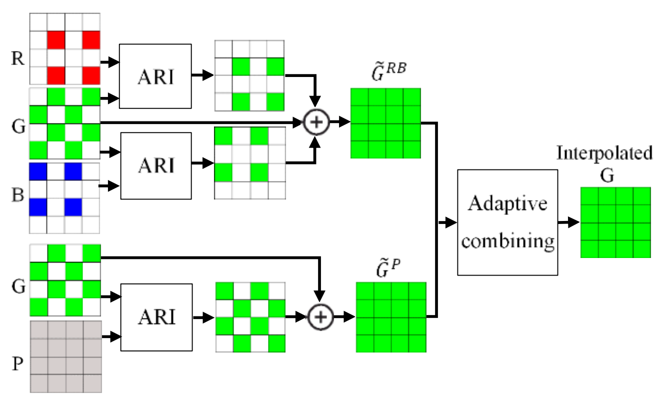


Figure 3. 4 The overall flow of CI applied to the green (G) channel.

Interpolating the G channel using the panchromatic image is quite like the one described in Section 3.2.2. The only difference is that the LRP image is not sparse; i.e. all pixel values are available; hence no interpolation is needed for LRP and it is fixed throughout the process. In other words, at each iteration, the LRP image is directly used as the guidance to generate the estimates of the G pixels and is not updated. For the adaptive combining step in Figure 3.4, two images ($\tilde{G}^{RB}$ and $\tilde{G}^{P}$) with the related minimum smoothed criterion images ($S^{\tilde{G}^{RB}}, S^{\tilde{G}^{P}}$), obtained from equations (3.20) to (3.22), are available. Finally, the interpolated G channel of the LRC image, $\tilde{G}$, is obtained through the weighted averaging as:

$$\tilde{G} = \frac{w_1.\tilde{G}^{RB} + w_2.\tilde{G}^{P}}{w_1 + w_2}, \; w_1 = \frac{1}{S^{\tilde{G}^{RB}}}, \;\; w_2 = \frac{1}{S^{\tilde{G}^{P}}}. \tag{3.25}$$

3.2.3.2 Interpolation of the R and B Channels

So far, two interpolated images, LRP and G, are available to guide R and B interpolations. The procedure we suggest here is different than the one in (Monno et al., 2017). We propose to use an iterative version of MLRI. Figure 3.5 illustrates the overall flow of CI for the R channel. The B interpolation is performed in the same manner. The procedure involves two interpolations, and for each, one of the full images (P or G) is used as the guidance. Therefore, two images are obtained ($\tilde{R}^{G}$ and $\tilde{R}^{P}$) which are then combined to generate the interpolated R image.

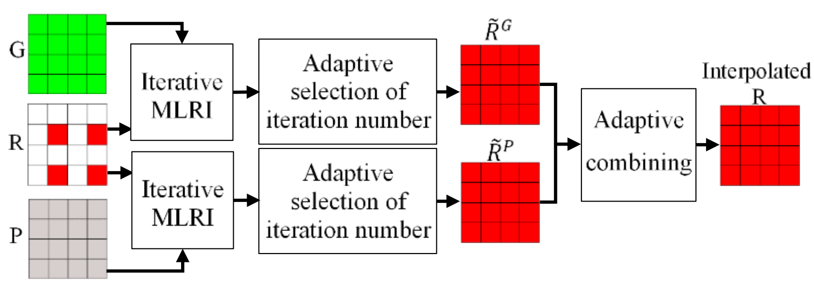


Figure 3. 5 The overall flow of CI applied to the red (R) channel. The blue interpolation is performed in the same manner

The main difference between iterative MLRI and iterative directional MLRI (Section 3.2.2.1) is that iterative MLRI is not applied directionally anymore. Besides, P and G channels are fixed throughout the procedure and are not updated at each iteration; thus, the difference between the results of each iteration is due to the GF window resizing. This section explains how to obtain $\tilde{R}^P$. The same procedure can be followed to obtain $\tilde{R}^G$. The estimation of the R pixels ($\hat{R}$) at each iteration are calculated using:

$$\hat{R}_k(i,j) = a_k^R(i,j).P(i,j) + b_k^R(i,j). \tag{3.26}$$

The coefficient, a^R, is then calculated in local windows using the GF procedure and by minimizing the following cost function:

$$a_{\omega,k}^R = arg\,min \sum_{i,j\in\omega} \left(\tilde{\nabla}^2 R_k\,(i,j) - \tilde{\nabla}^2 \hat{R}_k\,(i,j)\right)^2$$

$$= arg\,min \sum_{i,j\in\omega} \left(\tilde{\nabla}^2 R_k\,(i,j) - a_{\omega,k}^R\,.\tilde{\nabla}^2 P\ \ (i,j)\right)^2, \tag{3.27}$$

where $\tilde{\nabla}^2(.)$ is the approximated Laplacian operator, and $\tilde{\nabla}^2\hat{R}$ is obtained by convolving $\hat{R}$ with a 2-dimensional sparse Laplacian kernel as:

$$\tilde{\nabla}^2 R_k = \tilde{R}_{k-1} * L, \quad L = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}. \tag{3.28}$$

The coefficient b^R is obtained using equation (3.9). The final values of a^R and b^R, obtained from different sliding windows are calculated using equation (3.10), and the estimates are calculated using equation (3.26). The residuals are then obtained by subtracting the estimates from the originally available values in the input R:

$$r_k^R(i,j) = R(i,j) - \hat{R}_k\,(i,j),\ \ (i,j) \in \chi^R, \tag{3.29}$$

in which χ^R is the set of positions where the R values are available. The residuals for the missing pixels are calculated using a bicubic (BC) interpolation of the obtained residuals:

$$r_k^R(s,t) = r_k^R(i,j) * BC,\ \ (i,j) \in \chi^R \ and \ (s,t) \in \chi^{R'}, \tag{3.30}$$

where $\chi^{R'}$ is the set of positions where the R values are missing. Finally, the results are added to the estimates to generate the interpolated image of each iteration:

$$\tilde{R}_k(i,j) = r_k^R(i,j) + \hat{R}_k(i,j),\ \ (i,j) \in \chi^{R'}. \tag{3.31}$$

For the next iteration, the window size in GF is enlarged and a new set of a^R and b^R, is calculated. We initially set the window size as 5×5 which is enlarged by 1 for height and by 2 for width at each iteration. The iteration continues until the maximum iteration number (11 in our experiment) is reached.

For the adaptive selection of the iteration number here, we need to modify the equations (3.20) to (3.22). There are no directional results here and the LRP and G images are fixed throughout the procedure. The differences between the estimated pixels at k-th iteration and the previous interpolated pixels are calculated using:

$$d_k\ (i,j) = \hat{R}_k(i,j) - \tilde{R}_{k-1}(i,j) \tag{3.32}$$

Then, the criterion image at k-th iteration is obtained through:

$$crt_k(i,j) = \left(d_k(i,j)\right)^2 . \delta d_k(i,j),$$

$$\delta d_k(i,j) = |d_k\ (i-1,j) - d_k\ (i+1,j)| + |d_k\ (i,j-1) - d_k\ (i,j+1)| \tag{3.33}$$

Finally, the smoothed criterion image is obtained using:

$$S_k(i,j) = \ g(crt_k(i,j). \tag{3.34}$$

For each pixel, the minimum value among the smoothed criterion images is considered as the optimum value. For combining the obtained images, $\tilde{R}^G$ and $\tilde{R}^P$ in Figure 3.5, two optimum smoothed criterion images are available based on equations (3.32) to (3.34). Finally, the interpolated R/B channel of the LRC image is obtained through the weighted averaging in the same manner as equation (3.25).

3.2.4 Phase (iii): Fusion

So far, through the mentioned procedure, two images are generated: one LRC image at half resolution level and one HRP image at the high resolution level. Besides, Figure 3.1 (c) shows that the spectral range of the panchromatic sensor covers the spectral range of all RGB channels. This motivates the idea of incorporating pan-sharpening techniques from which Remote Sensing (RS) applications have benefited a lot over the last two decades. Pan-sharpening involves fusing an HRP image with an LRC image to generate a high resolution colour image, whose spectral information comes from the LRC image, and the sensitivity/spatial resolution is provided by the HRP image.

Based on Figure 3.1 (c), different colour channels contribute differently to generating the spectral response of the panchromatic sensor. This is neglected in equation (3.5) by dedicating equal weights for the three R, G, and B channels. However, such an approximation is acceptable in that step since equation (3.5) mostly supports the spatial relationship between panchromatic and RGB values which is more vital for restoring the spatial details of the HRP image. The spectral relationship is considered in the fusion procedure by calculating the contribution of each colour channel.

3.3 Experimental Results and Discussion

The evaluation routine in the literature is to simulate CFA2.0 images. However, simulation cannot fully reflect the nature of real imaging. This is especially true for panchromatic pixels as they are simulated by averaging R, G and B values. Figure 3.1 (c) shows that averaging the colour values is not necessarily equivalent to what a real panchromatic sensor senses. Thus, the evaluation using real CFA2.0 images is more favourable. Accordingly, we designed two scenarios; 1) simulated CFA2.0 images using the Kodak benchmark dataset, and 2) real CFA2.0 images captured by a camera equipped with a Truesense sensor.

For the fusion step, we have applied the UNB-Pansharp method which has been widely used in satellite images (Y. Zhang & Mishra, 2014) and has been studied for UAV (Jabari et al., 2017) and terrestrial applications (Khiabani, 2015). In this method, the least squares technique is used to establish the best possible relationship between the colour and panchromatic channels (Y. Zhang & Mishra, 2014). Thus, the procedure of this method is well aligned with the idea of considering the contribution of colour channels to the panchromatic spectral response. Also, the high speed of the UNB-Pansharp method is a great help to make our multi-step procedure operational for real-world applications. It is noteworthy that other fusion methods could be used, however, the comparison of different fusion methods and their effects on the quality of the demosaicked image is not in the scope of this study.

3.3.1 Simulated CFA2.0 Images

In this section, the Kodak dataset[2] which consists of 24 full resolution images of size 3072×2048 was used. These images were sampled based on CFA2.0 by keeping one out of the R, G, and B values for colour pixels, and by averaging the R, G, and B values for the panchromatic pixels. Then, different demosaicking methods were applied. Finally, the objective comparison was performed by measuring two mostly used quality metrics. This procedure was performed in two noise-free and noisy cases. For the noisy case, the images are corrupted with additive Gaussian white noise of standard deviation 5.

To see if the collaborative interpolation has any benefit over the independent interpolation, we have also applied II+fusion of (Kwan et al., 2017). However, to have a fair comparison, we had to implement other fusion and demosaicking methods of what was studied in (Kwan et al., 2017). More specifically, we used the same UNB-Pansharp method to fix the fusion step; and we applied the standard ARI method for the demosaicking step. Therefore, comparing the results of II+fusion with CI+fusion can lead us to the comparison of collaborative interpolation with the independent one. We also have compared our method to Condat's method with two regularizations, Tikhonov (Tik; Condat, 2009), and total variation (TV) minimization (Condat, 2014, 2012). In Condat's method, the regularization parameter (known as μ) has a great impact on the results. We empirically selected μ as 0.5 and 0.6 for Tik and TV methods respectivly, in a way that it

[2]Available at: https://www.math.purdue.edu/~lucier/PHOTO_CD/D65_TIFF_IMAGES/

minimizes the mean square error using the available low resolution version of the Kodak dataset[3] (with images of size 768×512).

For objective comparisons, the colour peak signal-to-noise ratio (CPSNR) and the CIE-LAB error are measured. The CPSNR is based on PSNR which is derived from the mean square error (MSE) between the reference (I) and the demosaicked ($\widehat{I}$) images as (Losson et al., 2010):

$$MSE^k = \frac{1}{3MN} \sum_{k=R,G,B} \sum_m \sum_n (I^k(m,n) - \hat{I}^k(m,n))^2,$$

$$CPSNR = 10 \log_{10}(\frac{255^2}{MSE^k}), \tag{3.35}$$

where M and N denote the dimension of each colour channel. CPSNR is measured in dB and the higher the CPSNR value, the higher the demosaicking quality.

To estimate the colour error between any two images, it is recommended to use the CIE-LAB colour space, since this space is close to a perceptually uniform color space (Losson et al., 2010). The colour error is then estimated by measuring the Euclidian distance between two colour vectors in the CIE-lab space. To do so, first the colour from RGB space is transferred into CIE-XYZ space using a linear transformation. Then the components in CIE-LAB space are calculated using (Losson et al., 2010):

$$L^* = \begin{cases} 116 \times f(Y/Y_n) - 16 & if\ Y/Y_n > 0.008856 \\ 903.3 \times (Y/Y_n) & otherwise \end{cases}$$

$$A^* = 500 \times (f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right))$$

$$B^* = 200 \times (f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right))$$

[3] Available at: http://r0k.us/graphics/kodak/

$$f(x) = \begin{cases} \sqrt[3]{x} & if\ Y/Y_n > 0.008856 \\ 7.787x + (16/116) & otherwise \end{cases} \tag{3.36}$$

where (X_n, Y_n, Z_n) is the reference white colour vector. Then, the error caused by demosaicking is the average of the spectral distances calculated for all the pixels as (Losson et al., 2010):

$$\Delta E^{L^*A^*B^*}(I, \hat{I}) = \frac{1}{MN} \sum_m \sum_n \sqrt{\sum_{k=L^*,A^*,B^*} (I^k(m,n) - \hat{I}^k(m,n))^2}\ . \tag{3.37}$$

The lower the CIE-LAB error, the higher the quality of the demosaicked image.

Table 3.1 reports the measured CPSNR and CIE-LAB error. For the noise-free case, our method outperforms II+fusion by 0.69dB in CPSNR, and by 0.09 in CIE-LAB error, and outperforms Condat’s method by 1.20dB and 1.40dB in CPSNR, and 0.05 and 0.20 in CIE LAB error, respectively for TV and Tik regularizations. The proposed method shows its capability more in the noisy case, where it outperforms II+fusion by 1.93dB in CPSNR, and by 0.43 in CIE-LAB error, and outperforms Condat’s method by 2.23dB and 1.92dB in CPSNR, and 0.78 and 0.76 in CIE LAB error, respectively for TV and Tik regularizations. For the noisy case, the proposed method performs the best for all the images in CPSNR, and the best for most (almost 90%) of the images in CIE-LAB error. Thus, given that the panchromatic image is interpolated with high quality, using its pixels in both demosaicking and fusion steps is even more efficient in the noisy case.

Interestingly, the methods involving fusion in their procedures (CI+fusion and II+fusion), perform better than other two methods in the presence of the noise (considering metrics in Table 3.1). However, comparing these two methods in Table 3.1, given that the fusion step is fixed in both, shows that CI performs better than II.

Table 3. 1 Demosaicking quality comparison using simulated CFA2.0 images.

	Image	CPSNR				CIE-LAB Error			
		Condat-Tik	Condat-TV	II+fusion	CI+fusion	Condat-Tik	Condat-TV	II+fusion	CI+fusion
Noise free	1	42.80	**44.27**	42.50	43.63	3.23	**1.24**	1.43	1.33
	2	37.40	41.49	43.76	**44.32**	1.39	1.86	1.27	**1.21**
	3	41.92	36.83	45.10	**45.42**	1.78	1.41	1.04	**1.01**
	4	41.65	45.62	45.07	**45.63**	1.41	1.04	1.10	**1.03**
	5	43.06	43.27	44.30	**45.08**	1.28	1.18	1.25	**1.16**
	6	43.40	**45.00**	43.37	44.77	1.28	**1.08**	1.25	1.14
	7	43.71	44.48	44.64	**44.97**	2.53	**1.06**	1.17	1.14
	8	38.05	**38.33**	36.60	37.38	**1.73**	2.48	2.81	2.66
	9	41.56	**41.64**	39.65	39.80	1.56	**1.71**	2.06	2.01
	10	**42.51**	42.50	40.63	40.81	**1.25**	1.55	1.82	1.78
	11	43.71	44.07	43.42	**44.21**	**1.08**	1.14	1.30	1.22
	12	44.91	**46.76**	44.86	45.68	1.94	0.86	1.02	**0.98**
	13	**40.11**	39.98	38.57	39.78	**1.56**	1.86	2.16	1.94
	14	40.24	35.67	43.73	**44.41**	1.37	1.86	1.18	**1.13**
	15	42.32	45.20	44.10	**45.52**	1.08	0.99	1.07	**0.98**
	16	45.30	**45.89**	44.25	45.30	1.77	**1.00**	1.19	1.10
	17	**41.55**	41.09	40.02	40.28	2.53	**1.77**	2.02	1.93
	18	**38.88**	38.62	37.70	38.11	**1.57**	2.49	2.86	2.70
	19	42.42	42.71	42.24	**42.98**	1.87	**1.44**	1.63	1.51
	20	40.15	38.57	39.61	**40.22**	1.73	**1.81**	1.98	1.87
	21	41.67	**42.54**	40.65	41.30	1.71	**1.57**	1.86	1.74
	22	41.42	39.78	41.96	**42.63**	1.55	1.60	1.62	**1.51**
	23	40.94	30.61	45.50	**45.91**	1.09	2.14	1.01	**0.97**
	24	44.07	43.84	43.78	**44.23**	1.49	1.03	1.12	**1.02**
	Avg.	41.82	41.62	42.33	**43.02**	1.66	1.51	1.55	**1.46**
Noisy (standard deviation=5)	1	36.37	36.64	37.41	**38.43**	2.88	2.81	2.30	**2.11**
	2	34.44	36.46	37.78	**38.67**	3.99	2.76	2.08	**1.94**
	3	36.15	34.04	32.23	**36.85**	2.86	3.17	3.72	**2.45**
	4	36.09	36.85	37.04	**38.77**	3.04	2.72	2.37	**1.99**
	5	36.43	36.42	36.05	**38.79**	2.94	2.93	2.60	**2.07**
	6	36.50	36.77	38.27	**38.74**	2.72	2.69	1.98	**1.94**
	7	36.56	36.68	36.53	**39.04**	2.77	2.77	2.37	**1.88**
	8	34.75	34.84	34.54	**35.56**	3.60	3.61	3.36	**3.15**
	9	36.06	36.06	35.60	**37.34**	2.97	3.03	2.73	2.40
	10	36.31	36.31	37.21	**37.72**	2.89	2.96	2.38	2.30
	11	36.57	36.61	37.47	**38.88**	2.80	2.84	2.23	**1.96**
	12	36.77	37.04	37.45	**39.31**	2.53	2.51	2.03	**1.72**
	13	35.60	35.50	35.29	**36.89**	3.22	3.25	2.99	**2.62**
	14	35.64	33.30	34.33	**37.81**	3.08	3.59	3.03	**2.19**
	15	36.27	36.76	36.15	**39.00**	2.85	2.75	2.48	**1.90**
	16	36.83	36.92	37.88	**39.13**	2.71	2.76	2.12	**1.88**
	17	36.07	35.93	36.49	**37.53**	3.10	3.19	2.63	2.46
	18	35.13	34.97	34.60	**35.79**	3.67	3.73	3.60	**3.28**
	19	36.29	36.36	36.60	**38.27**	2.93	2.94	2.60	**2.22**
	20	35.62	34.99	35.69	**37.21**	3.03	3.09	2.67	**2.39**
	21	36.09	36.30	35.48	**37.87**	3.02	3.00	2.97	**2.33**
	22	36.01	35.41	35.41	**37.97**	3.03	3.12	2.90	**2.31**
	23	35.83	29.79	31.83	**36.65**	2.91	3.74	3.93	**2.56**
	24	36.65	36.60	37.42	**38.86**	2.75	2.81	2.16	**1.94**
	Avg.	36.04	35.73	36.03	**37.96**	3.01	3.03	2.68	**2.25**

Figure 3.6 compares parts of the results for image 5 of the Kodak dataset in both noise-free and noisy cases. The original image is shown at the top, while the next two rows contain the results for the noise-free and noisy cases, respectively. The images show that the proposed method gives the least artifacts and colour noise and better preserves the spatial details of the original images in both noise-free and noisy cases.

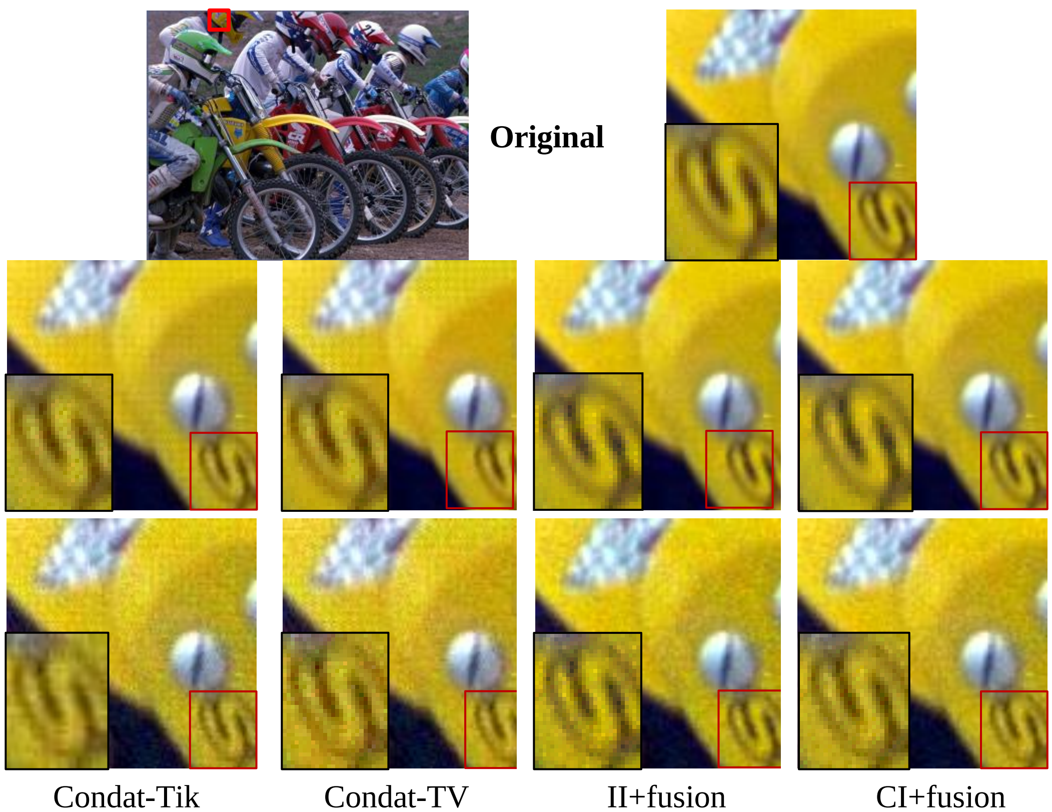


Figure 3. 6 Visual comparison of the demosaicked images for the noise-free case (first row) and the noisy case (second row) using image 5 of the Kodak dataset. The parts in the red boxes are enlarged for a better comparison.

3.3.2 Real CFA2.0 Images

The camera used in this experiment is a BOBCAT IGV-B2320 v1.0, manufactured by IMPERX, which is built around the KAI-04050C sensor with a Truesense sparse CFA. For the specifications of the camera see (*Bobcat Intelligent camera series IGV-B2320*, 2013). In addition to the methods mentioned in the previous section, we compare our results to the demosaicked images provided by the IMPREX toolkit (Bobcat GEV) coming with the camera. The images were taken in different imaging conditions using a Kowa LM12HC lens with a focal length of 12.5 mm. The aperture was set to F/8 and the exposure time was set to 60ms. Table 3.2 summarizes the imaging specifications. Image 1 represents a noise-free condition, where the light source (MK-II Fiber optic direct current) was set at maximum intensity. While images 2-4 represent noisy conditions, where the light source intensity was decreased by one order of magnitude.

Table 3. 2 The imaging specifications.

Image	Measured illuminance [lux]	White-balance Coefficients [R, G, B]	Analog Gain	Description
1	81.7	[256, 345, 700]	0	No optical density (OD) filter was used for the light source. Auto white-balance adjustment was performed.
2	8.89	[256, 345, 700]	0	O.D. Filter was changed.
3	8.89	[256, 385, 820]	512	Analog gain adjustment and auto white-balance adjustment were performed.
4	8.89	[256, 385, 820]	0	The analog gain was reset to 0, but white-balance coefficients were kept.

Figure 3.7 shows the indoor scene used for imaging with the standard charts designed for quality measurement. The reference colour object, the X-Rite colour checker

chart ('Area1' in Figure 3.7) was used for measuring CIE-LAB error. The accurate reference colour values for this chart are already made available by the producer. Therefore, the results with less colour distance from the reference values have higher colour quality. We also measured the SNR using four colour patches of the colour checker chart shown in 'Area 2' of Figure 3.7. Since there is no reference image in this dataset, we calculate SNR as the ratio between the mean value and the standard deviation of the brightness values in each patch and then we averaged the values over the four patches of Area 2. With such a calculation, we can interpret SNR as a meaningful metric for sensitivity.

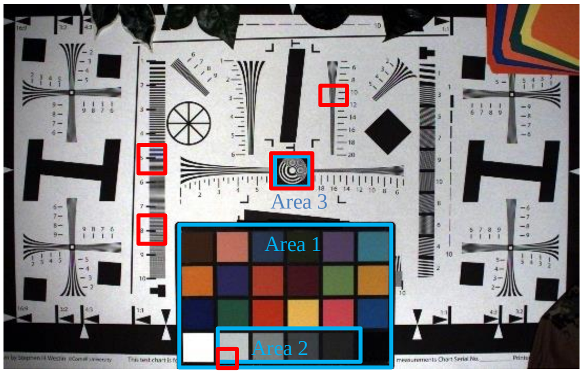


Figure 3. 7 The indoor scene with standard charts used for real CFA2.0 image capturing. The areas in blue boxes, Area 1, 2 and 3, are used to measure CIE-LAB error, SNR, and channel discrepancy (CD), respectively. The areas within red boxes are enlarged for visual comparisons in Figure 3.9.

The mentioned metrics do not reflect a comprehensive interpretation of image quality because they are limited to Area 1 and 2, where colour artifacts do not normally appear due to the lack of high frequency information (e.g. edges, textures, etc.). To consider colour artifacts in our comparisons, we designed and measured a metric called 'Channel Discrepancy' (CD) in an appropriate area ('Area 3' in Figure 3.7). CD is based on this assumption that the R, G and B pixels in Area 3 should have the same value since the scene contains only black and white circles. However, this does not happen due to the generated colour artifacts. Therefore, if CD represents the difference between the maximum and the minimum of R, G, B values at pixel (i, j),

$$CD(i,j) = \max\{R(i,j), G(i,j), B(i,j)\} - \min\{R(i,j), G(i,j), B(i,j)\}, \qquad (3.38)$$

then the average of CD over the selected area can be a good parameter to reflect the severity of colour artifacts. The lower the average value of CD, the higher the quality of the demosaicked image. It is noteworthy that CD is also affected by noise; however, we selected Area3 which is more affected by colour artifacts rather than by noise.

Table 3.3 reports the measured metrics for different images. Considering the colour quality, the CIE-LAB error caused by our method in images 1 and 3 is close to the best one, belonging to the camera's results, and in images 2 and 4, is less than the one caused by other methods. Comparing the SNR values, our proposed method has resulted in higher values in all images except for 1, in which it produces the closest value to the best one, belonging to the camera's result. This again shows that incorporating high quality panchromatic pixels in both demosaicking and fusion steps leads to less noise which is reflected in higher SNR values, especially in noisy conditions. Moreover, our proposed method shows its strength better when we compare the generated colour artifacts in different results. The CD values for the images produced by other methods are higher than the ones generated by our method. This shows that the colour artifacts produced by other methods are significantly higher than the ones generated by our method (this is further proved by the visual comparison of the results provided in Figure 3.9). Here, it is again observed that the methods which involve fusion in their procedures (CI+fusion and

II+fusion) perform better than other methods. However, the superiority of CI+fusion especially in SNR and CD values shows the better performance of the collaborative interpolation compared to the independent interpolation in retrieving the spatial details.

Table 3. 3 Demosaicking quality comparison for real CFA2.0 images.

		Condat-Tik	Condat-TV	II+fusion	Camera's result	CI+fusion
CIE-LAB error	1	18.46	19.03	17.63	**16.16**	17.96
	2	26.04	26.81	25.74	28.97	**25.73**
	3	19.14	18.76	18.19	**16.88**	17.31
	4	26.22	26.99	25.79	28.23	**25.72**
Average SNR	1	40.31	39.97	53.35	**68.93**	54.62
	2	15.12	11.02	24.96	13.90	**25.20**
	3	12.06	11.65	25.69	25.10	**26.09**
	4	12.09	10.77	24.03	14.01	**24.37**
Average CD	1	17.07	15.90	5.05	21.46	**4.77**
	2	22.31	21.94	4.99	33.16	**4.94**
	3	19.10	18.41	5.04	22.87	**4.77**
	4	20.00	19.32	4.83	30.61	**4.72**

For the visual evaluations, we first provide the outputs of collaborative and independent interpolations, i.e. the HRP and LRC images, only for image 2 (due to space limitations) in Figure 3.8. The HRP image obtained from CI (top right image) provides much higher spatial details than the one obtained without involving colour pixels (top left image). This proves that incorporating colour pixels in panchromatic interpolation helps in reconstructing the HRP image with high spatial fidelity. Besides, comparing the LRC image obtained from CI (bottom right image) with the one obtained independently of panchromatic pixels (bottom left image) shows that involving panchromatic pixels in demosaicking the colour image results in less artificial colour.

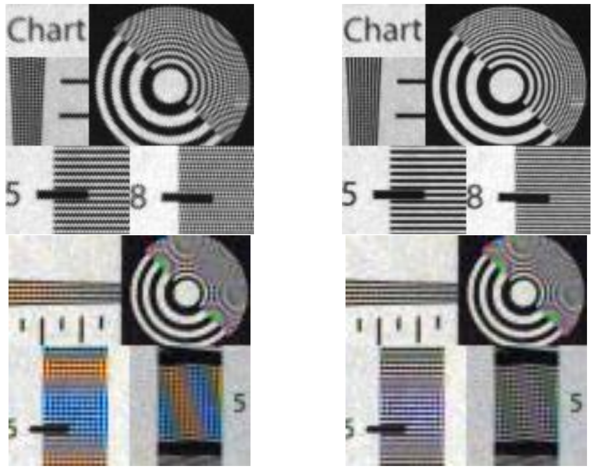
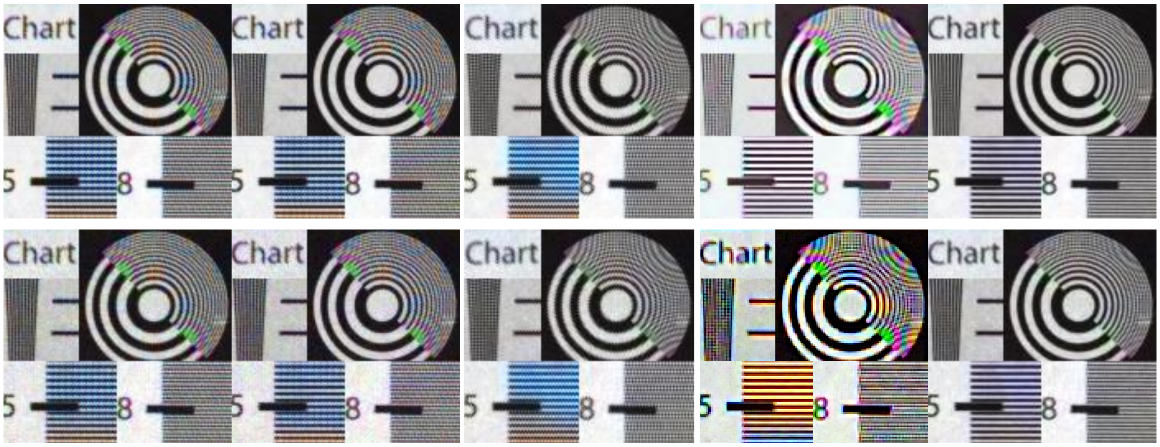


Figure 3. 8 Visual comparison of the HRP (first row) and LRC (second row) images obtained through independent (left column) and collaborative (right column) interpolations for image 2. Different areas are cropped and enlarged for a better comparison of the differences.

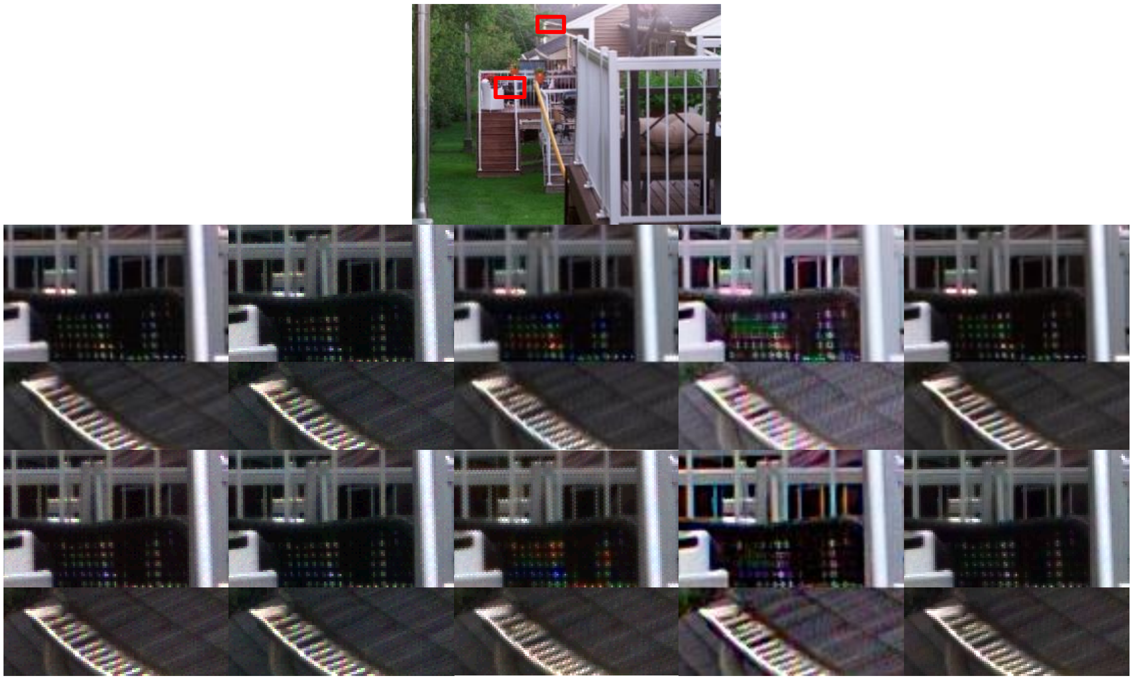
Figure 3.9 compares the results for two images (1 and 2). As can be seen, our method results in far fewer colour artifacts, especially in the areas related to high-frequency information. It produces the least jagged edges (zipper effect) and colour noise and retrieves the spatial resolution (fine details) more precisely (circles, horizontal and vertical lines are better distinguishable). This is while all other results suffer from severe colour artifacts and jagged edges, and the fine details are lost.



Condat-Tik Condat-TV II+fusion Camera’s results CI+fusion

Figure 3. 9 Visual comparison of the results for image 1 (first row) and image 2 (second row) captured in high light and low light conditions, respectively. For the imaging specifications of images 1 and 2 refer to Table 3.2

To complete our experiment, we took images of an outdoor scene (top image in Figure 3.10) in the natural evening light. Since no standard charts were available, we provide only the visual comparison of the results in Figure 3.10. To involve different noise levels, we simulated different lighting conditions by manipulating the aperture., for high lighting (noise-free) and low lighting (noisy) conditions, the aperture was set to F/5.6 and F/8, respectively. For both conditions, the exposure time was set to 60ms, and the auto white- balance adjustment was performed without applying analogue/digital gain. In different parts of the images, our results have far fewer colour artifacts and higher object details (the texture of the chair and the edge of the roof in Figure 3.10). Besides, comparing Figures 3.9 and 3.10 with Figure 3.6, the differences in the efficiency of the methods in real images are much more noticeable than the simulated images. Based on our experiments, methods may perform differently in real CFA2.0 images and this indicates the need to test any RGBW method on the real datasets.



Condat-Tik Condat-TV II+fusion Camera's results CI+fusion

Figure 3. 10 Visual comparison of the results for an outdoor scene (top image), captured in high light (second row) and low light (third row) conditions. The areas within the red boxes are enlarged.

Overall, quantitative and qualitative evaluations, in both simulated and real datasets, show the improved performance of the proposed method compared to the other investigated methods as well as the camera's results. This improvement is especially noticeable in retrieving fine details of the scene objects without injecting high colour artifacts or noise. This can be explained by the fact that in our proposed method, the features of the panchromatic image are injected into the colour image in 2 steps: once in half resolution level through demosaicking the extracted Bayer image and once in full resolution level through the fusion procedure. Also, evaluations indicate that the strategy of the two-step information injection is even more efficient in the tested noisy cases (based on Table 3.1 and the noisy images 2-4 in Table 3.3).

3.4 Conclusions

RGBW colour filter arrays contain panchromatic (or white) pixels along with Red, Green, and, Blue pixels to increase the sensitivity of the sensors. Kodak released sensors with a specific RGBW CFA called Truesense or CFA2.0, which are available in the market. However, reconstructing a full resolution colour image from CFA2.0 with acceptable quality is still challenging.

In this study, we proposed a procedure for CFA2.0 demosaicking. This procedure involves a collaborative interpolation (CI) between panchromatic and colour pixels and a fusion method to fuse the interpolated panchromatic and colour images. In this procedure, the features of the panchromatic pixels (i.e. high sensitivity and high spatial details) are injected into the colour image in two resolution levels. This greatly improved the quality of the reconstructed full resolution colour images.

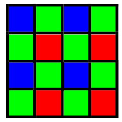
To evaluate the performance of the proposed procedure simulated CFA2.0 images, obtained from the Kodak benchmark dataset, and real CFA2.0 images, captured by a camera kit equipped with Truesense sparse CFA, were used. The evaluations demonstrate that our proposed procedure significantly reduces the colour artifacts and zipper effects, produces less colour noise, and retrieves finer spatial details compared to the other investigated methods as well as the camera's results. Therefore, the proposed demosaicking method has great potential to improve the quality of the images obtained from Truesense sensors in real-world applications.

4 Deep Joint Demosaicking and Denoising for RGBW Colour Filter Arrays

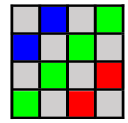
4.1 Introduction

Most digital cameras today use only one sensor chip instead of three, to reduce manufacturing costs, which gives a one-channel raw image. A raw image is a mosaic image with a pattern of alternating Red (R), Green (G), and Blue (B) colour pixels captured using a Colour Filter Array (CFA). A CFA is a matrix of filter elements placed on top of the sensor. Each filter element passes only a specific wavelength of red, green, or blue light. Thus, at each pixel of the raw image, only one of the three R, G, or B values is recorded. Since, filtering reduces the light energy, the sensitivity of the sensor decreases, resulting in a low signal-to-noise ratio (SNR), especially in low light or noisy conditions. To improve the sensitivity, panchromatic (P) or white (W) filters, transparent to the visible range of the electromagnetic spectrum, are incorporated into RGB CFAs to generate RGBW CFAs (Hamilton & Compton, 2007; Kijima et al., 2006). Kodak was the first to build an RGBW-based sensor, called CFA2.0 with a specific pattern (Figure 4.1) called Truesense sparse pattern (AND9180/D-TRUESENSE Sparse Color Filter Pattern, 2014).

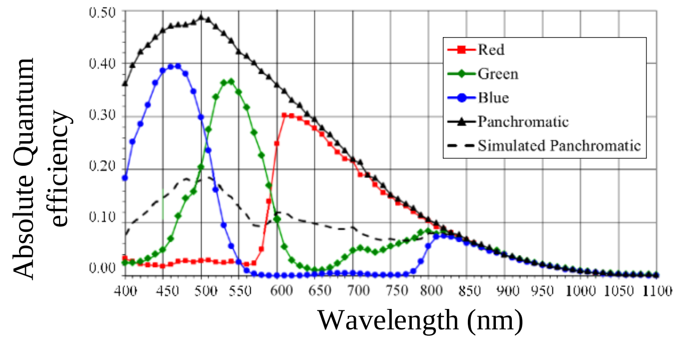
To convert a one-channel raw image into a three-channel full resolution colour image, a processing pipeline is employed with the first two crucial steps being demosaicking and denoising. Demosaicking refers to the processing through which the missing pixel values are recovered. While demosaicking for RGB CFAs has been



(a) RGB CFA (Bayer)



(b) RGBW CFA or CFA2.0 (Truesense)



(c) Quantum efficiency of Kodak Truesense sensor (plot adopted from (AND9180/D-TRUESENSE Sparse Color Filter Pattern, 2014)). The simulated panchromatic plot is obtained by averaging Red, Green, and Blue plots.

Figure 4. 1 Comparing RGB and RGBW CFAs and the related spectral responses.

extensively studied (Din et al., 2020; Elgendy et al., 2021; Khadidos et al., 2021; Losson et al., 2010; Monno et al., 2017; Park & Jeong, 2019), there are few papers related to RGBW demosaicking.

Kodak proposed a two-resolution-level method involving two main steps, demosaicking panchromatic and colour images and integrating the obtained images using a very simple fusion method (Hamilton & Compton, 2007). The same demosaicking-and-fusion process was used in other studies with different demosaicking and fusion techniques. In (Kwan et al., 2017), the panchromatic and colour images were demosaicked independently using NAT (L. Zhang et al., 2011) and Malvar-He-Cutler (Malvar et al., 2004) algorithms respectively. For the fusion step, they recommended using the pan-sharpening methods utilized in Remote Sensing. Later, Kwan and Chou (Kwan & Chou, 2019) used a pre-trained deep network (Gharbi et al., 2016) for demosaicking the colour and panchromatic image. The colour image is demosaicked independently of the panchromatic image, and for demosaicking the panchromatic image, only red and blue pixels of the demosaicked colour image were used. Finally, the demosaicked images were fused using a pan-sharpening method called hybrid colour mapping (Kwan & Chou, 2019). In another study (Chung et al., 2020), the panchromatic image was first interpolated using a cross shape-based colour difference approach. For the fusion, they followed the approach in (Hamilton & Compton, 2007). Then, the obtained colour image was further enhanced through two more steps; an iterative error compensation-based approach and a refinement process based on the RGBW raw image. Few other methods perform RGBW demosaicking in one resolution level by estimating the luminance and chrominance components. The estimation is performed either in the frequency domain (Rafinazari & Dubois, 2016; J. Wang et al., 2011), or in the spatial domain through the so-called universal methods (Condat, 2009, 2012; Kang & Jung, 2020; Singh & Singh, 2012; C. Zhang et al., 2016).

Although it has been theoretically demonstrated that integrating panchromatic pixels into RGB CFAs boosts the sensor's SNR, RGBW CFAs have not been widely used due to the inefficiency of the demosaicking methods offered thus far. Besides the fact that limited studies have focused on RGBW demosaicking, denoising is usually neglected in such studies. None of the mentioned studies has performed a comprehensive investigation about the effect of different noise levels (usually one noise level is investigated) on the results.

Demosaicking by itself is considered an ill-posed problem, where two-thirds of the data is missing. With noise intervention, this problem gets more complicated since three colour values are to be found from a single noisy observation. Apart from the ill-posed nature of both demosaicking and denoising steps, the order of these two steps in the pipeline is another issue to be investigated. Fortunately, the dilemma of demosaicking first or denoising has been investigated for the RGB CFAs (Jin et al., 2020; Kokkinos & Lefkimmiatis, 2019). Studies show the superiority of a joint demosaicking and denoising procedure over the approaches that break the problem into sequential steps (Jin et al., 2020; Kokkinos & Lefkimmiatis, 2019). Moreover, the joint demosaicking and denoising methods are mostly based on deep learning, especially deep Convolutional Neural Networks (CNN; Din et al., 2020; N. Dong et al., 2021; W. Dong et al., 2018; Gharbi et al., 2016; Kokkinos & Lefkimmiatis, 2019; Pan et al., 2020; Syu et al., 2018; H. Tan et al., 2022; Wu et al., 2021). These networks are mainly designed based on the idea of the residual networks (ResNet) presented by (He et al., 2016). The popularity of the residual networks stems from the fact that they enable the creation of very deep networks by addressing the notorious vanishing gradient problem. In fact, they are aimed to learn some residuals defined simply as the subtraction of a layer's output from its input. This is done using the shortcut or skip connections that skip one or more layers by directly connecting the input of one layer to the input of the next layer or a few layers after. It has been demonstrated that training this type of network is simpler than training normal deep CNNs (He et al., 2016).

The achievements of the RGB demosaicking studies motivate us to bring such concepts into RGBW demosaicking by aiming at a joint demosaicking and denoising approach via designing a deep CNN model based on residual learning. However, the main issue we counter when designing a learning-based algorithm for RGBW demosaicking is the shortage of ground-truth images, more specifically, the lack of panchromatic images captured simultaneously with the existing colour images. The non-learning-based methods, such as (Condat, 2009, 2012; Kwan et al., 2017; Rafinazari, 2017; Singh & Singh, 2012; J. Wang et al., 2011; C. Zhang et al., 2016), used benchmark datasets (mainly Kodak and IMAX), for their evaluations by simulating the panchromatic pixel as the average of the R, G, and B pixels. If we use such a simulation to obtain a training dataset, the network would learn during the training procedure that the panchromatic pixel values are the average of the colour pixel values. Such a trained network will not function effectively when exposed to a real RGBW image, captured by an RGBW sensor, where the relationship between panchromatic and colour pixels is not necessarily a simple average. Figure 4.1 (c), better shows that averaging the colour pixels is not what a real panchromatic sensor senses. Thus, one of the important challenges when implementing deep learning methods for RGBW demosaicking is to generate more realistic panchromatic images to be used in the training dataset.

In this study, inspired by the state-of-the-art demosaicking methods for RGB CFAs, we propose a Deep CNN for a Joint Demosaicking and Denoising of the RGBW images, and we call it RGBW-DJDD (or briefly RGBW-DJ). It is worth noting that the way we integrate deep learning into RGBW demosaicking is different than (Kwan & Chou, 2019). In (Kwan & Chou, 2019), the RGBW raw image is somehow reshaped into an RGB raw image, for which the pre-trained network for RGB CFAs (Gharbi et al., 2016) is used. In contrast, our network is directly applicable to RGBW raw images. In other words, in (Kwan & Chou, 2019), the network is neither designed nor trained for the RGBW pattern. However, our approach focuses on designing, training, and testing a deep network specifically for RGBW raw images. The proposed design is inspired by the network in (Syu et al., 2018), called Demosaicking CNN_ Vey Deep (or DMCNN_VD) for RGB CFAs.. The DMCNN_VD design uses only one skip connection between the input and the last block. However, in RGBW-DJDD, inspired by ResNet (He et al., 2016), each consecutive two blocks are connected with a skip connection. We also have brought the idea of employing a mosaic layer proposed in (Fletcher, 2018) into our design, which is responsible for learning the colour filter pattern. Generally, the proposed network comprises a mosaic layer, an initial and a final layer, and 10 repeating blocks between them. Each block contains a convolutional layer, a batch normalization layer, and an activation layer. Since the network is meant to be used for real CFA2.0 sensors, the necessary factors for practical use, such as generality, speed, and low computational load are considered in its design. For generality, the mosaic layer helps the network to be reusable for other RGBW patterns as well, although our experiment here is limited to a specific sensor with a known pattern. To provide the speed and low computational load, we designed a lightweight network with 10 blocks to reduce the trainable parameters. Moreover, we avoid using multiple models as used in (Gharbi et al., 2016; D. S. Tan et al., 2018) to reduce the cost of multiple training and we avoid initial demosaicking as in (D. S. Tan et al., 2018) or downsampling as in (Gharbi et al., 2016). Also, unlike (Gharbi et al., 2016), our model does not require the noise level as input and can handle denoising with unknown noise level (i.e., blind denoising).

To solve the problem of the shortage of ground-truth panchromatic images, we propose to simulate them using the widely available RGB datasets. To do so, a linear transformation is established to convert the RGB images into simulated panchromatic images. This is inspired by a similar idea in other applications such as pan-sharpening, where the panchromatic image is spectrally related to the colour channels through a linear combination (Fathollahi & Zhang, 2020; Meng et al., 2019). When the panchromatic images are simulated, we can generate full resolution 4-channel images containing the already existing RGB channels plus the simulated panchromatic channel. These images are then used as the training and testing datasets to train and test RGBW-DJDD.

To evaluate the proposed RGBW demosaicking method, unlike most of the papers, we do not limit our experiments to only benchmark datasets (mainly IMAX and Kodak). As a further step, we have used real RGBW raw images captured by BOBCAT IGV-B2320 camera which is equipped with a Truesense sensor. We have compared our model to Kodak's technique (Hamilton & Compton, 2007), the variational approach VAR (Condat, 2009), Effective Three-Stage (ETS) method (Chung et al., 2020), and Kwan's method (Kwan & Chou, 2019). We refer to Kwan’s method as ‘RGB-DJ+fusion’, since it comprises two basic processes: an RGB deep joint demosaicking and denoising (briefly RGB-DJ) from (Gharbi et al., 2016), and a fusion step. For real CFA2.0 images, we also have compared our results to the camera's software outputs. Our model achieved higher quality based on quantitative metrics, colour peak signal to noise ratios (CPSNR) and CIE-LAB error, as well as visual comparisons. Evaluations using real RGBW raw images captured in different imaging settings and corrupted with real imaging noise instead of the simulated one, further prove the generality of our network. It is noteworthy that since the main focus of this research is to improve the quality of the RGBW demosaicking results, comparing the performance of the RGB and RGBW CFAs is outside the scope of this research and will be followed in the future.

The rest of the chapter is organized as follows. Section 4.2 first gives a detailed description of the proposed network. The results are presented and discussed in Section 4.3. Finally, Section 4.4 concludes the chapter.

4.2 RGBW Deep Joint Demosaicking and Denoising

4.2.1 Architecture

The architecture of the proposed RGBW-DJDD is based on DMCNN_VD (Syu et al., 2018) which is itself inspired by another reference archetcture in (J. Kim et al., 2016). The DMCNN_VD network consists of 20 layers with the kernel size of 3, where each layer is composed of a convolution layer, a batch normalization layer and a SELU activation layer. It uses only one skip connection between the first and the last block to incorporate residual learning into their design. However, in the spirit of ResNet (He et al., 2016), we propose that instead of using only one skip connection between the input and the last block as in DMCNN_VD, we employ skip connections between the subsequent two blocks. We also have incorporated the concept of a mosaic layer, as presented in (Fletcher, 2018), into our design. In addition, since the mentioned networks were all originally developed for RGB CFAs, we had to modify their design to make them RGBW compatible. The modification of the design includes the number of blocks, the size of the convolutional kernels, and the patch sizes. The rest of this section goes into the proposed network design in further depth. As depicted in Figure 4.2, the RGBW-DJDD network generally consists of one mosaic layer (M), followed by an initial convolutional layer ($\mathrm{Conv_I}$), N repeating blocks, each of two connected with a skip connection, and one final convolutional layer ($\mathrm{Conv_F}$).

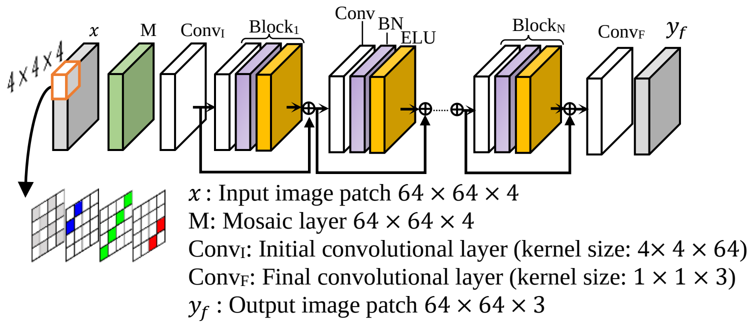


Figure 4. 2 The architecture of the proposed RGBW-DJDD network. It consists of 10 blocks with the residual learning strategy. Each block is composed of a convolution layer, a batch normalization (BN) layer and an ELU activation layer.

The mosaic layer is responsible to learn the pattern of the CFA through the weights assigned to each pixel. In fact, the weights to be learned, determine whether a specific wavelength is allowed to pass or not. To do so, the raw image is rearranged into a four-channel image by leaving the three missing channels blank for each pixel. Setting the patch size as 64×64, the input of the network would have the size of 64×64×4. The mosaic layer, then, consists of 64×64×4 weights, all initialized by 0.5. In the mosaic layer, a raw mosaic image is generated by performing a per-pixel dot product between the input patch, x, and the weights of the mosaic layer, W_M, as:

$$y = \sum_{i=1}^{4} x_i . W_{M_i}. \tag{4.1}$$

This simulates how a raw mosaic image is generated through a CFA that blocks certain wavelengths and passes others. Then, the weights are multiplied by the simulated raw image from the first step, y, resulting in a four-channel image. In other words, the i-th channel of the mosaic layer's output is obtained using

$$y_{M_i} = y \times W_{M_i}. \tag{4.2}$$

Finally, the CFA weights are updated by comparing the predicted image with the reference ground-truth image. The initial convolution layer is intended to collect low resolution local features using 64 4×4 kernels. The kernel weights are initialized based on Lecun_normal initializer (LeCun et al., 1998), which draws samples from a truncated normal distribution centred on 0 with the standard deviation of $(\frac{1}{number\ of\ weights})^{0.5}$. If W_0 and B_0 represent the parameters (kernel weights and the biases) of this layer and $*$ shows the convolution operator, the output of the layer is

$$\mathrm{y}_0 = W_0 * y_M + B_0. \tag{4.3}$$

Each block consists of 64 4×4 convolutional filters, initialized based on Lecun_normal initializer (LeCun et al., 1998), followed by a batch normalization layer, and an ELU (Exponential Linear Unit; Clevert et al., 2016) activation layer. Based on ResNet (He et al., 2016), we can write the output of the n-th block as:

$$y_n = \mathcal{F}_n(y_{n-1}, \{W_n\}, B_n) + y_{n-1}. \tag{4.4}$$

Where, $y_1 = \mathcal{F}_1(x, \{W_n\}, B_n) + y_0$, and $\mathcal{F}$ represents the residual mapping function for which the weights W_n and the biases B_n should be learned:

$$\mathcal{F}_n = W_n * \delta(\sigma(\mathcal{F}_{n-1})) + B_n, \quad \mathcal{F}_1 = W_1 * x + B_1, \tag{4.5}$$

with δ and σ representing the batch normalization and ELU (Clevert et al., 2016) functions, respectively. The last 1×1 convolutional layer is intended to reduce the number of channels to 3 for the output RGB image. In other words, the colour is reconstructed using local details from a collection of features.

4.2.2 Joint Denoising

Based on (K. Zhang et al., 2017) the residual learning strategy helps to gradually separate the image structures from noise through intermediate blocks. Thus, the architecture discussed in the previous section is already capable of denoising by integrating residual learning into CNN. The only thing left is to expose the network to noisy images throughout the training procedure. The network would learn to suppress noise by comparing its noisy output against the noise-free ground-truth image.

Rather than having numerous networks specialized for different noise levels, it is preferable to have a single network that can cope with images contaminated with various noise levels. To do so, the images used for training should be randomly contaminated with different noise levels. According to (Gharbi et al., 2016), noise in the white-balanced gamma-corrected images can be represented as signal-independent white Gaussian noise. Thus, based on (Gharbi et al., 2016), each image is corrupted with a centred additive Gaussian noise of standard deviation σ (known as noise level), which is randomly picked from a range $[\sigma_1, \sigma_2]$ before training. However, unlike (Gharbi et al., 2016), we do not explicitly introduce the noise level to the network. This leads us to blind denoising, with the expectation that the network would perform better for noise levels within the range with which it has been trained, rather than for noise levels beyond that range.

4.2.3 Preparing Training Dataset

Figure 4.3 shortly reviews the procedure of preparing the mosaic images required to train the network. As mentioned, the main challenge of integrating deep learning into

RGBW demosaicking is the lack of full-resolution 4-channel images including panchromatic and RGB channels. Inspired by the idea used mostly in pan-sharpening techniques (Fathollahi & Zhang, 2020; Meng et al., 2019), we propose to simulate the panchromatic image using available 3-channel datasets and through a linear transformation applied to the R, G, and B channels. To do so, for each RGB image, I_{RGB}, three random positive weights (w_R, w_G, and w_B) are selected in a way that satisfies $w_R + w_G + w_B = 1$. The simulated panchromatic image, P_s, is then obtained through:

$$P_s = w_R * R + w_G * G + w_B * B. \tag{4.6}$$

Each weight belongs to the range [0,1]. The network can choose 0 as the weight for each of the colour channels; however, as it learns, it will be less likely to choose zero weights for various channels. When P_s is generated, it is appended, as the fourth channel, to the originally available RGB channels, which gives our required image, I_{RGBP}.

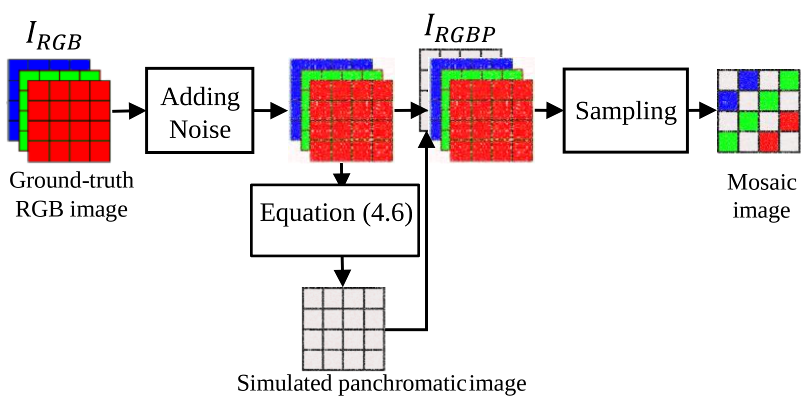


Figure 4. 3 Preparing the mosaic images required to train RGBW-DJDD.

Considering equation (4.6) as the weighted average of the RGB channels and considering that averaging is one of the means for noise suppression, it can be guaranteed that the simulated panchromatic image through equation (4.6) would have a higher SNR than other RGB channels. Thus, to provide higher SNR for the simulated panchromatic image, the process of adding noise, described in Section 4.2.2, should be performed on the original RGB image (I_{RGB}), before generating P_s. When I_{RGBP} is generated, the mosaic image is created by sampling only one colour channel per pixel according to the RGBW pattern.

4.2.4 Training

For training, the output patch from the network (y_{f_i}) is compared to the ground-truth image patch (x_i), and the kernel weights and biases are updated by minimizing the mean absolute error between the two patches. In other words, we define the loss function as:

$$L(\{W,B\}_j) = \frac{1}{n}\sum_{i=1}^{n}\left|y_{f_i} - x_i\right|, \tag{4.7}$$

for the j-th layer, and n is the number of training patches. The optimization is performed using nADAM (Dozat, 2016), a stochastic gradient descent method. We use a batch size of 500, and a learning rate of 10−3. All other parameters are left at the value recommended by (Dozat, 2016). The number of trainable parameters in total is 530,179. The implementation is carried out by support provided by Compute Canada on Cedar Cluster using one node with 128GB RAM and one GPU (NVIDIA P100-PCIE-12GB).

The RGB dataset we used to train our network is the MS COCO 2014 dataset (Lin et al., 2014), which contains 3174 images. We additionally augmented the training patches by scaling, randomly rotating in 90 degrees steps, and randomly flipping. This gives a tenfold increase in the training data as well as the rotational and translational invariance. For evaluation of the training, we used the Waterloo Exploration Dataset (WED) with 4744 images (Ma et al., 2017). The training ran for 3000 epochs and took about 36 hours.

4.3 Experiments and Results

We evaluate our network (briefly referred to as RGBW-DJ in this section) using two kinds of datasets. First, we follow the routine in the literature by simulating CFA2.0 images using benchmarks. For this case, we compare our algorithm against other RGBW demosaicking approaches, including Kodak's method (Hamilton & Compton, 2007), VAR (Condat, 2009), ETS (Chung et al., 2020), and RGB-DJ+fusion (Kwan & Chou, 2019). Second, to evaluate the efficiency of the network for real-world applications, we used real CFA2.0 images captured by a camera equipped with a Truesense sensor. For this case, we excluded the ETS method from our evaluations since we could not obtain satisfactory results from the provided code. Instead, we compared our results with the outputs of the camera's software.

4.3.1 Simulated CFA2.0 Images

The Kodak and the IMAX (also known as McMaster) datasets are the most often used demosaicking benchmarks. These datasets are full-resolution 3-channel images, which are sampled based on any desired CFA/CFA2.0 patterns in different studies. In our study, we follow the procedure shown in Figure 4.3 to simulate the CFA2.0 pattern. The outputs of any demosaicking method are then compared to the original full-resolution images by measuring two mostly used quality metrics, the colour peak signal-to-noise ratio (CPSNR) and the CIE-LAB error (Losson et al., 2010).

For the joint demosaicking and denoising, as stated in Section 4.2.2, we train the network on images corrupted with different noise levels randomly selected from the range of [0, 20]. For evaluation, we used images distorted at 5 distinct noise levels,

$\sigma_{noise} = \{0, 5, 10, 15, 20\}$, with $\sigma_{noise} = 0$ denoting the noise free case. Tables 4.1 and 4.2 report the measured CPSNR and CIE-LAB error for different methods in two cases, $\sigma = 0$ and $\sigma = 20$, on IMAX and Kodak benchmarks, respectively. These two metrics were calculated in the same manner describe in Chapter 3, using equations (3.35) to (3.37). Instead of providing the results for all the images in all 5 noise levels, we averaged CPSNR and CIE-LAB error values over the images in each dataset, which are reported in Table 4.3.

Our method generates superior results quantitatively in both benchmarks. Especially, our model provides significantly higher quality compared to the other DL-based algorithm, RGB-DJ+fusion. It is mainly because our technique employs all pixels, regardless of whether they are panchromatic or RGB, in the interpolation of missing pixels. Unlike in RGB-DJ+fusion, where the colour pixels are demosaicked first without the inclusion of the panchromatic pixels, and the panchromatic pixels are demosaicked independently of the green pixels. The performance of our network is also noticeable in noisy cases. When compared to other approaches, our model produces the greatest quality results from the lowest noise level to the highest.

Table 4. 1 Demosaicking performance comparison on IMAX dataset.

Noise level	Image	Kodak's method		VAR		ETS		RGB-DJ+fusion		RGBW-DJ	
		CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB
Noise free	1	24.27	9.60	25.98	8.19	26.03	7.73	25.08	8.24	**28.69**	**5.88**
	2	28.74	5.71	32.25	4.43	32.62	4.19	29.94	4.91	**34.48**	**3.55**
	3	24.67	8.54	29.94	5.58	30.88	5.01	25.84	6.97	**33.66**	**3.78**
	4	25.11	6.72	31.74	3.97	34.00	3.02	28.17	4.42	**36.68**	**2.29**
	5	28.43	5.82	30.25	4.92	30.56	4.50	29.70	4.81	**33.47**	**3.54**
	6	30.70	5.18	31.77	4.66	32.97	4.04	32.37	4.14	**34.91**	**3.48**
	7	29.36	5.10	34.73	3.35	35.88	3.14	30.21	4.37	**38.90**	**2.08**
	8	28.86	4.11	34.51	2.63	35.13	2.49	30.64	3.23	**38.67**	**1.92**
	9	29.16	5.00	32.62	3.87	33.43	3.43	31.06	3.88	**36.54**	**2.71**
	10	31.85	3.94	34.63	3.13	35.12	2.89	33.44	3.12	**37.64**	**2.41**
	11	33.02	3.51	35.41	2.97	36.26	2.63	34.61	2.80	**38.39**	**2.25**
	12	30.80	3.67	35.12	2.68	36.11	2.44	32.15	3.05	**37.99**	**2.24**
	13	34.14	2.77	37.49	2.23	38.49	2.09	36.05	2.34	**40.02**	**2.04**
	14	33.66	3.11	36.38	2.58	37.00	2.44	34.64	2.79	**38.08**	**2.32**
	15	33.44	3.09	36.14	2.57	36.25	2.48	34.69	2.69	**38.05**	**2.16**
	16	27.22	6.73	29.24	5.19	30.11	4.67	28.96	5.19	**32.77**	**3.86**
	17	27.58	7.71	28.15	7.26	28.48	6.49	28.46	6.21	**31.51**	**4.81**
	18	26.85	6.47	31.90	4.43	33.15	3.83	29.54	4.73	**35.65**	**3.07**
$\sigma_{noise} = 20$	1	22.38	13.06	22.79	12.44	21.96	12.25	22.26	13.43	**25.67**	**8.47**
	2	24.91	10.17	25.03	9.61	23.80	9.58	24.68	10.86	**29.17**	**5.96**
	3	22.61	12.69	24.19	10.77	23.14	10.31	22.81	12.54	**27.91**	**6.33**
	4	22.98	11.09	24.84	9.25	23.78	8.45	23.88	10.31	**30.13**	**4.43**
	5	24.54	10.73	24.22	10.49	22.99	10.28	24.21	11.34	**28.48**	**6.21**
	6	25.32	10.41	24.56	10.46	23.33	10.41	24.73	11.62	**28.84**	**6.82**
	7	24.97	10.03	25.13	9.32	23.88	8.93	24.93	10.04	**29.71**	**4.08**
	8	25.14	8.68	25.79	7.93	24.72	7.60	25.31	9.29	**31.53**	**3.39**
	9	24.96	10.02	24.90	9.62	23.65	9.54	24.80	10.43	**30.18**	**5.26**
	10	25.97	8.73	25.45	8.56	24.13	8.74	25.47	9.48	**30.48**	**5.02**
	11	26.38	8.42	25.67	8.55	24.43	8.83	25.82	9.48	**31.04**	**4.82**
	12	25.41	9.32	25.08	9.02	23.75	9.10	25.27	9.39	**31.21**	**4.31**
	13	25.99	9.09	25.07	9.09	23.64	9.19	25.88	9.14	**32.77**	**3.78**
	14	26.31	8.64	25.52	8.64	24.17	8.70	25.91	9.41	**31.60**	**4.51**
	15	26.44	8.01	25.75	7.99	24.41	8.23	26.03	8.75	**31.73**	**4.16**
	16	24.25	10.28	24.31	9.78	23.48	9.76	24.04	11.13	**27.54**	**6.92**
	17	24.42	11.23	23.97	11.41	22.99	11.54	23.82	12.15	**26.88**	**8.41**
	18	23.92	10.55	24.85	9.37	23.77	9.11	24.39	10.41	**29.39**	**5.36**

Table 4. 2 Demosaicking performance comparison on Kodak dataset.

Noise level	Image	Kodak's method		VAR		ETS		RGB-DJ+fusion		RGBW-DJ	
		CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB	CPSNR	CIE-LAB
Noise free	1	25.52	7.52	32.13	4.54	34.44	3.61	26.78	6.29	**37.8**	**2.51**
	2	31.40	3.46	32.73	4.90	36.52	2.21	31.70	3.19	**39.0**	**1.99**
	3	32.74	2.43	36.34	2.21	40.58	**1.38**	33.74	2.24	**41.3**	1.44
	4	30.91	3.49	35.34	3.12	37.97	2.14	32.54	3.04	**39.5**	**1.86**
	5	25.51	7.76	31.21	4.83	33.03	4.07	26.52	6.53	**36.6**	**2.80**
	6	26.82	5.63	33.30	3.53	36.00	2.73	28.15	4.90	**38.2**	**2.34**
	7	30.38	3.68	36.31	2.43	37.97	1.90	32.76	2.77	**40.2**	**1.76**
	8	22.85	9.44	29.93	5.20	32.90	3.92	24.80	7.50	**36.3**	**2.76**
	9	29.92	3.30	37.02	2.12	39.42	1.78	32.74	2.73	**41.0**	**1.55**
	10	29.44	3.39	36.55	2.16	38.37	1.86	32.47	2.79	**40.6**	**1.62**
	11	28.28	4.82	34.20	3.01	36.27	2.56	29.35	4.23	**38.7**	**2.02**
	12	29.67	3.02	37.82	2.07	37.23	**1.66**	33.35	2.56	**40.8**	**1.66**
	13	23.49	9.40	29.18	6.09	30.87	5.48	24.01	8.58	**34.3**	**3.73**
	14	27.85	5.78	31.63	4.21	33.73	3.34	28.50	5.08	**36.7**	**2.51**
	15	28.92	3.71	34.71	2.83	36.90	2.20	31.35	3.22	**38.7**	**1.96**
	16	31.19	3.67	37.05	2.38	39.62	1.94	31.43	3.41	**41.4**	**1.64**
	17	30.21	3.58	36.67	2.19	38.45	1.96	32.13	2.93	**40.3**	**1.61**
	18	27.29	5.95	32.26	4.08	33.82	3.65	27.72	5.41	**35.8**	**2.89**
	19	27.30	5.08	33.95	3.12	37.44	2.42	29.11	4.22	**39.0**	**2.03**
	20	29.44	3.33	35.60	2.36	37.26	1.93	31.15	2.93	**39.3**	**1.90**
	21	27.56	5.22	33.36	3.50	35.22	2.95	28.83	4.47	**38.3**	**2.15**
	22	29.44	4.55	34.25	3.34	35.50	2.88	30.19	4.18	**37.3**	**2.56**
	23	32.19	2.42	35.25	2.63	39.12	**1.56**	34.13	2.20	**40.4**	1.64
	24	26.29	5.97	31.42	3.85	32.87	3.40	26.59	5.29	**34.9**	**2.72**
$\sigma_{noise} = 20$	1	23.06	12.02	24.52	10.24	23.40	9.72	23.38	11.08	**28.31**	**5.25**
	2	25.59	8.73	25.12	8.49	23.81	8.68	25.41	8.65	**30.83**	**4.31**
	3	25.79	9.23	25.18	9.20	23.79	9.21	25.85	9.13	**32.32**	**3.64**
	4	25.31	9.54	25.07	9.20	23.68	9.24	25.52	9.09	**30.74**	**4.39**
	5	23.10	12.42	24.49	10.74	23.44	10.23	23.16	12.00	**28.40**	**5.91**
	6	23.76	10.81	24.78	9.65	23.61	9.30	24.15	10.19	**29.34**	**4.79**
	7	25.17	9.90	25.09	9.42	23.68	9.26	25.39	9.54	**31.43**	**4.45**
	8	21.39	13.50	24.12	10.70	23.38	9.79	22.15	12.15	**28.24**	**5.49**
	9	24.96	9.63	25.02	9.23	23.64	9.04	25.49	9.60	**32.16**	**3.66**
	10	24.80	9.76	24.99	9.39	23.62	9.16	25.46	9.42	**31.95**	**3.91**
	11	24.48	10.45	24.94	9.71	23.69	9.40	24.62	10.36	**30.09**	**4.45**
	12	24.93	9.09	25.17	8.73	23.76	8.67	25.76	8.66	**31.83**	**3.83**
	13	21.81	13.32	23.88	11.28	23.10	10.70	22.00	12.57	**26.67**	**6.46**
	14	24.29	10.98	24.62	10.24	23.47	9.79	24.17	10.80	**28.91**	**5.19**
	15	24.87	9.69	25.30	9.09	24.05	9.11	25.45	9.55	**31.06**	**4.33**
	16	25.39	9.97	25.05	9.64	23.68	9.36	25.32	9.70	**31.11**	**3.97**
	17	25.30	9.83	25.30	9.39	23.95	9.11	25.58	9.62	**31.63**	**3.79**
	18	24.06	11.36	24.71	10.49	23.56	10.25	23.96	11.32	**28.63**	**5.58**
	19	24.02	10.76	24.84	9.72	23.64	9.35	24.48	10.27	**30.45**	**4.28**
	20	25.21	8.60	25.69	8.04	24.58	7.91	25.57	8.60	**31.72**	**3.64**
	21	24.12	10.81	24.76	9.88	23.53	9.56	24.44	10.19	**29.87**	**4.38**
	22	24.83	10.27	24.85	9.74	23.52	9.60	24.86	10.20	**29.71**	**5.09**
	23	25.66	9.05	25.13	8.98	23.73	9.04	25.80	9.05	**32.23**	**4.03**
	24	23.50	11.40	24.44	10.35	23.37	9.96	23.43	10.88	**28.76**	**5.09**

Table 4. 3 Demosaicking performance comparison on Kodak and IMAX datasets for noise free and all noisy cases. The values are averaged over the images in each dataset.

		$\sigma_{noise} =$	0	5	10	15	20
IMAX	Kodak's method	CPSNR	29.33	28.77	27.56	26.16	24.83
		CIE-LAB	5.38	5.97	7.13	8.55	10.06
	VAR	CPSNR	32.68	31.26	28.87	26.67	24.84
		CIE-LAB	4.15	4.92	6.30	7.90	9.57
	ETS	CPSNR	33.47	31.25	28.17	25.65	23.67
		CIE-LAB	3.75	4.64	6.14	7.80	9.48
	RGB-DJ + fusion	CPSNR	30.87	29.84	27.99	26.21	24.68
		CIE-LAB	4.33	5.31	6.93	8.71	10.51
	RGWB-DJ	CPSNR	**34.88**	**33.61**	**31.96**	**30.52**	**29.12**
		CIE-LAB	**3.46**	**3.88**	**4.51**	**5.20**	**5.93**
Kodak	Kodak's method	CPSNR	28.53	28.06	26.98	25.68	24.39
		CIE-LAB	4.86	5.64	7.07	8.73	10.46
	VAR	CPSNR	34.09	32.62	29.53	26.94	24.88
		CIE-LAB	3.36	4.07	5.80	7.71	9.65
	ETS	CPSNR	36.31	32.95	28.86	25.88	23.65
		CIE-LAB	2.65	3.81	5.61	7.52	9.39
	RGB-DJ + fusion	CPSNR	30.00	29.11	27.64	26.07	24.64
		CIE-LAB	4.19	5.05	6.64	8.39	10.11
	RGWB-DJ	CPSNR	**38.59**	**35.95**	**33.45**	**31.72**	**30.27**
		CIE-LAB	**2.31**	**2.82**	**3.49**	**4.17**	**4.92**

The improvement is also clearly visible in Figure 4.4 which shows the results for image 1 and image 19 of the IMAX and Kodak datasets, respectively. The original reference images are exhibited at the top, followed by the results in the next three rows for the noise free and two noisy cases, $\sigma_{noise} = 10$ and $\sigma_{noise} = 20$. The visual comparison of the results also confirms the quantitative evaluations. In noise free and noisy scenarios, the suggested technique creates the fewest artifacts and colour noise and better preserves the spatial details of the references. The improvement is most noticeable in noisier situations.

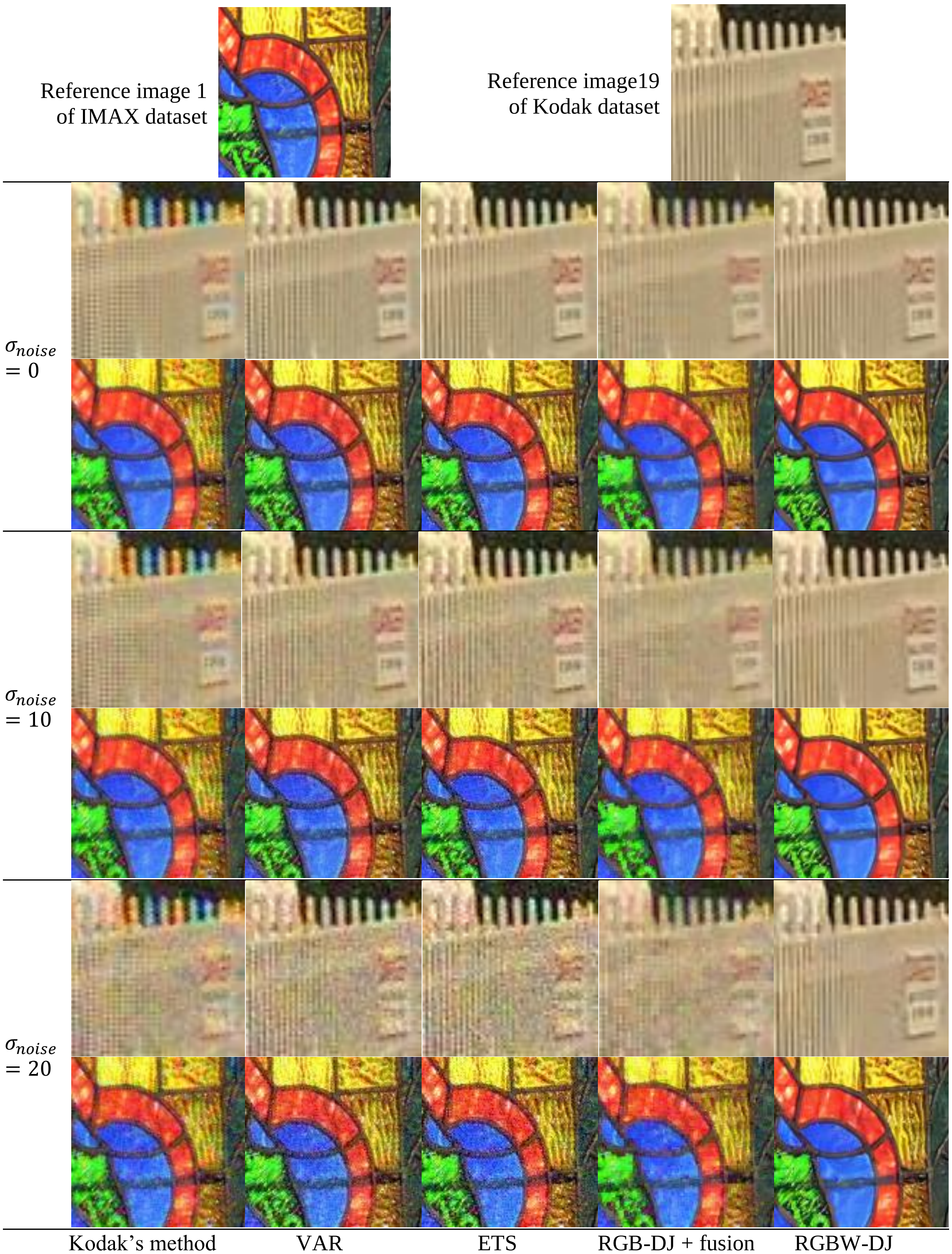


Figure 4. 4 Comparison of demosaicking results on 2 samples of IMAX and Kodak, under different noisy conditions.

To investigate the effect of the noise range on the training, we considered three scenarios, each time the network was trained with a different noise range: i) $\sigma_{noise} \in [0, 20]$, ii) $\sigma_{noise} \in [0, 10]$, and iii) we fixed the noise level as $\sigma_{noise} = 10$. Figure 4.5 shows the results averaged over both benchmarks to have an overall evaluation. Our results generally outperform other methods on all noise levels, regardless of the noise range with which the network is trained. Even with a smaller noise range $\sigma_{noise} \in [0, 10]$, the network performs better than other approaches when exposed to noise ranges outside the range with which it was trained. Moreover, as seen in Figure 4.5, training the network using a range of noise levels outperforms training it with only one noise level. When the network is exposed to different degrees of noise during training, the generality of the network is better provided. Considering all evaluated situations, the network trained with a broader noise range, $\sigma_{noise} \in [0, 20]$, performs better on average.

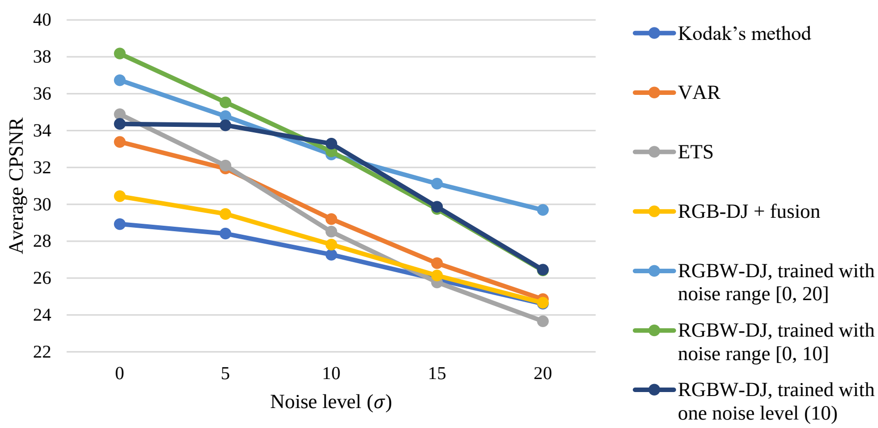


Figure 4. 5 Performance comparison of different methods at different Gaussian noise levels with standard deviation σ. The CPSNR is averaged over Kodak and IMAX images.

4.3.2 Real CFA2.0 Images

The camera utilized to capture real CFA2.0 images is an IMPERX BOBCAT IGV-B2320 v1.0, which is equipped with the KAI-04050C sensor with a Truesense sparse CFA (camera's specifications; *Bobcat Intelligent camera series IGV-B2320*, 2013). A Kowa LM12HC lens with a focal length of 12.5 mm was used and the images were captured under various imaging situations. The exposure length was set to 60ms, and the aperture was set to F/8. The imaging specifications are summarized in Table 4.4[4]. The light source (MK-II Fiber optic direct current) was adjusted to maximum intensity in Image 1 to approximate a noise free situation. Images 2-5, on the other hand, depict different noisy situations in which the light source intensity was reduced by one order of magnitude.

To be able to have quantitative evaluations, a scene with quality measurement charts was set up (Figure 4.6). The X-Rite colour checker chart ('Area1' in Figure 4.6) was utilized as the reference colour object for assessing CIE-LAB error. The manufacturer has already provided accurate reference colour values for this chart. Consequently, the images with the least colour distance from the reference values have superior colour quality.

---

[4] For this dataset, Canada's National Optics Institute (INO, https://www.ino.ca), set up the camera and the scene, equipped with the standard quality measurement charts, and took the images at the industry standard.

Table 4. 4 the Imaging Settings.

Image	Measured illuminance [lux]	White-balance Coefficients [R, G, B]	Analog Gain	Description
1	81.7	[256, 345, 700]	0	No optical density (OD) filter was used for the light source. Auto white-balance adjustment was performed.
2	8.89	[256, 345, 700]	0	Light source intensity was decreased by one order of magnitude. OD Filter was changed.
3	8.89	[256, 385, 820]	512	Analog gain adjustment and auto white-balance adjustment were performed.
4	0.74	[256, 345, 700]	0	Light source intensity was decreased by one order of magnitude. OD Filter was changed.
5	0.74	[256, 485, 963]	1023	Analog gain adjustment and auto white-balance adjustment were performed.

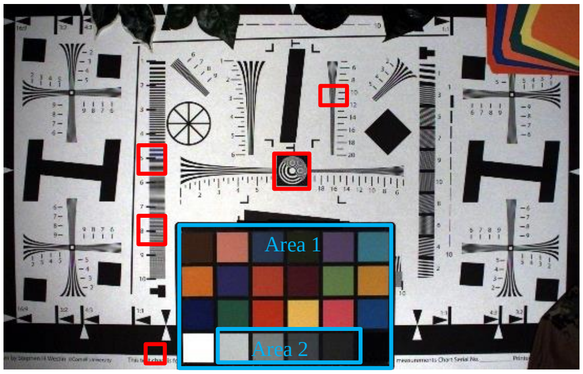


Figure 4. 6 The indoor scene with standard charts, used for real CFA2.0 image capturing. Area 1 and 2, are used to measure CIE-LAB error and SNR, respectively. The areas within red boxes are enlarged for visual comparisons in Figure 4.7.

We cannot compute CPSNR in the way we did for Kodak and IMAX since there is no reference image for this dataset. Instead, we calculated SNR as the ratio between the mean and the standard deviation of the pixel values in four patches, shown in 'Area 2' of

Figure 4.6, and then averaged the ratios over the four patches of Area 2. SNR measured in this manner can be taken as a sensitivity metric.

The nature of real imaging is always very different from the simulated situation, no matter how close that simulation is to reality. Accordingly, when it comes to real-world images, we still need to fine-tune the network to compensate for the difference between simulation and reality. Thus, for demosaicking real RGBW raw images, we fine-tuned our network using a dataset containing real full resolution panchromatic and colour images. These images were captured using two Huawei cell phones, P9[5] and Mate 9[6], which are equipped with dual cameras, i.e., two panchromatic and colour sensors. We also used a customized camera equipped with a couple of panchromatic and colour sensors, CIS-2521F sCMOS[7]. We took images from various indoor and outdoor scenes under different lighting conditions to obtain varying noise levels. Our fine-tuning dataset in total contains 170 images of various sizes (3900x2900, 3700x2700, 4900x3700 and 2360x1960). We additionally augmented the patches by scaling, rotating in 90 degrees steps, and flipping.

Here, we compare our results to the demosaicked images provided by the IMPREX toolkit (Bobcat GEV) that came with the camera, as well as the methods stated in the

[5] For the specifications see: https://www.gsmarena.com/huawei_p9-7972.php.

[6] For the specifications see: https://www.gsmarena.com/huawei_mate_9-8073.php

[7] For the specifications see: https://datasheets.globalspec.com/ds/2559/BAESystemsFairchildImaging/90479D4B-10C9-4465-93E7-BA0C85218E40

previous section. We exclude ETS since we could not obtain satisfactory results using the provided code. Table 4.5 reports the measured metrics for different images. Figure 4.7 compares the results for three images (1, 2, and 4 in Table 4.4).

Table 4. 5 Demosaicking performance comparison on real CFA2.0 images.

		Kodak's method	VAR	RGB-DJ + fusion	Camera's result	RGBW-DJ
Average SNR	1	30.72	40.31	58.19	68.93	**77.34**
	2	8.08	15.12	23.05	13.90	**38.68**
	3	13.23	12.06	24.07	25.10	**40.90**
	4	2.30	2.67	4.14	0.51	**7.44**
	5	2.70	2.29	4.48	4.07	**9.17**
CIE-LAB error	1	24.15	25.65	25.10	**16.16**	16.44
	2	27.61	26.04	25.47	28.97	**23.53**
	3	25.25	26.03	25.32	**16.88**	18.44
	4	38.38	37.66	35.38	56.16	**30.67**
	5	35.94	37.82	33.37	35.48	**28.58**

It should be noted that the camera's software outputs are not merely the result of demosaicking; the images are also subjected to further processing, i.e., colour correction. As a result, for the two images 1 and 3 in Table 4.5, the CIE-Lab error for the camera's results is less than that of all other methods followed very closely by our model at the second rank. Images 1 and 3 represent high light situations, one with the light intensity set to the highest level and the other with analogue gain applied (referring to Table 4.4). Despite the colour correction applied to the camera's results, our model is far superior to the Camera's result in CIE-Lab error under other noisier settings (images 2, 4, and 5). As can be observed in Figure 4.7, our approach creates significantly fewer colour artifacts, especially in areas with high-frequency information in various imaging situations. It produces the fewest jagged edges (zipper effect) and colour noise. In our results, circles, horizontal and vertical lines are more recognizable. Other outputs, on the other hand, have substantial colour artifacts and jagged edges, as well as lost fine details. Considering the results of RGB-DJ+fusion, they suffer from severe blurriness and all the fine details are lost (see the third column in Figure 4.7).

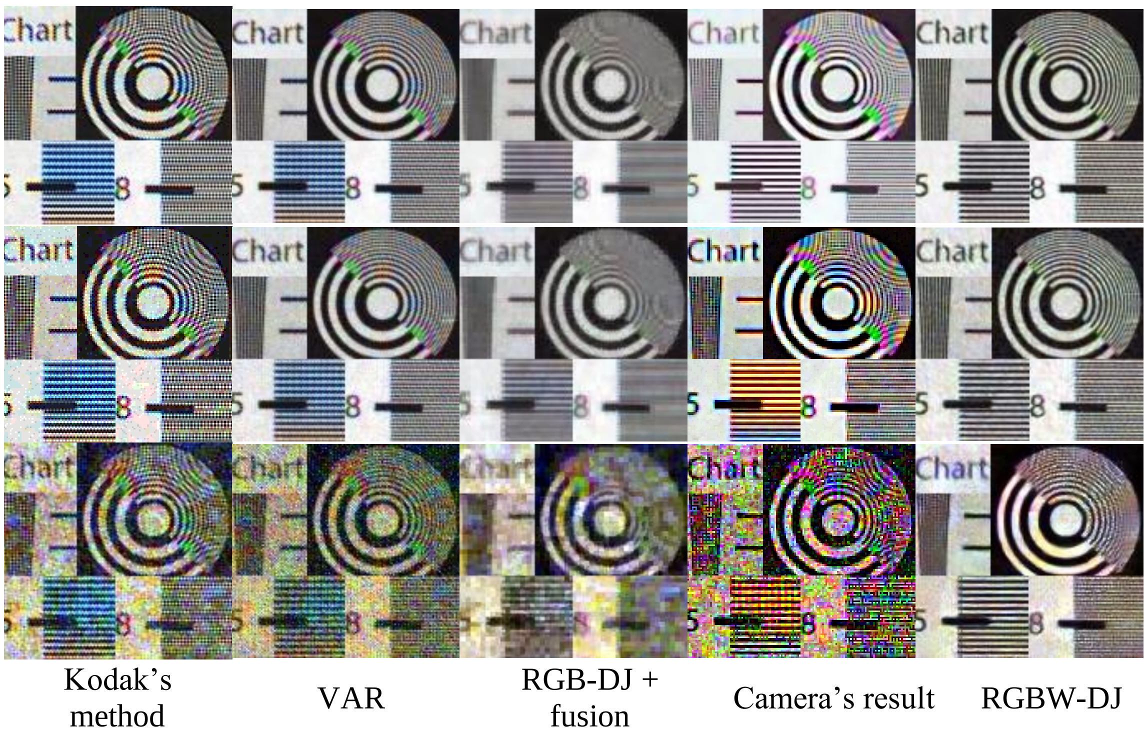


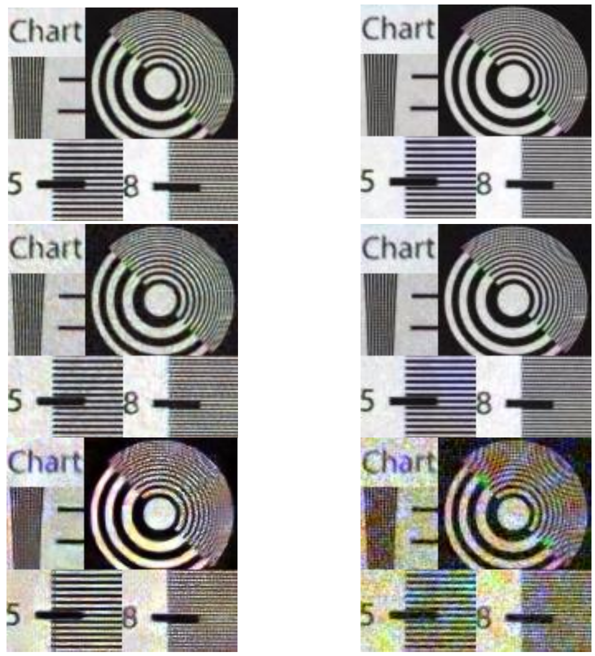
Figure 4. 7 Visual comparison of the demosaicking results for real CFA2.0 images; image 1 (first row), image 2 (second row), and image 4 (third row). The imaging settings are adjusted in such a manner that they imitate various noise situations (refer to Table 4.4).

It is worth mentioning that because the network was trained to decrease the white Gaussian noise, it's unlikely that any other noise type will be eliminated. Being at the early phases of using DL-based techniques to RGBW demosaicking, there is still room for improvements in complicated real-world noisy situations like the one illustrated in the last row of Figure 4.7.

4.3.2.1 Comparing CI+fusion and RGBW-DJ Methods

This section provides a brief comparison between the methods proposed in Chapters 3 and 4. Figure 4.8 provides a visual comparison of the results for images 1, 2 and 4 of Table 4.4. Comparing the results, it can be observed that in image 1 and 2, the spatial details of the lines are better preserved, and the text “Chart” has been less distorted in the images obtained by CI+fusion method. For image 4 in Figure 4.8, no meaningful differences in terms of retaining the fine spatial details can be observed, however, the colour noise in the RGBW-DJ results is significantly reduced when compared to the CI+fusion results. Comparing the average SNR values in Tables 4.5 and 3.3 for images 1, 2, and 3 shows that the RGB-DJ method has provided higher SNR values. This is because the RGBW-DJ model was trained to suppress the noise which results in greater SNR values compared to any method that does not involve denoising.

In summary, for noise-free and moderately noisy images, CI+fusion better maintains spatial features or gives higher spatial resolution, whereas RGBW-DJ delivers higher SNR. However, for images highly contaminated with noise, RGBW-DJ outperforms CI+fusion by better suppressing the colour noise.



RGBW-DJ CI+fusion

Figure 4. 8 Visual comparison of the proposed demosaicking methods in Chapters 3 and 4 for real CFA2.0 images; image 1 (first row), image 2 (second row), and image 4 (third row). The imaging settings are adjusted in such a manner that they imitate various noise situations (refer to Table 4.4).

4.4 Conclusions

To boost the sensitivity of the imaging sensors, RGBW colour filter arrays include panchromatic (or white) pixels alongside red, green, and blue pixels. Kodak was the first to develop an RGBW-based sensor, called CFA2.0 with a unique pattern called Truesense sparse pattern. Although panchromatic pixels give better SNR than RGB pixels, they pose issues in demosaicking, making RGBW CFAs unattractive in real-world applications.

In this study, we presented a deep CNN for RGBW joint demosaicking and denoising, RGBW-DJDD. Furthermore, to address the scarcity of ground-truth panchromatic images, we simulated them using a linear transformation with random weights applied to the widely available RGB datasets. We carried out a comparative study on simulated (using Kodak and IMAX benchmarks) and real CFA2.0 images. The findings showed that our model greatly eliminates colour artifacts and zipper effects, creates less colour noise, and recovers finer spatial features compared to the other evaluated methods as well as the camera's results. Therefore, the proposed method has the potential to put Truesense sensors to use in real-world applications.

The direction for future study would be to make the network robust against more complex types of noise that happen in real-world imaging. Besides, the focus of this research was on improving the performance of the RGBW demosaicking. Future studies can be persuaded by comparing RGB and RGBW CFAs through the performance comparison of the related demosaicking approaches at varying noise levels.

5 Summary and Conclusions

This chapter provides a summary of the research accomplished through the chapters 2 to 4 of this dissertation. It also summarizes the contributions and achievements of the research. Finally, it gives some suggestions and opens up directions for further research in this field.

5.1 Summary

This PhD dissertation proposed strategies to improve the image spatial quality in two different fields; the hyperspectral images in remote sensing, and the colour images obtained from RGBW CFAs in terrestrial applications. The spatial quality improvement is accomplished by using panchromatic sensors besides colour RGB sensors, which provide higher spatial quality than RGB sensors. The spatial quality enhancement is then followed in a computational processing manner through pan-sharpening in RS and RGBW demosaicking in terrestrial imaging. The second chapter in this dissertation addressed the problems associated with the hyperspectral pan-sharpening, while the third and fourth chapters involved the challenges and unaddressed topics in the field of RGBW demosaicking. More specifically,

- Chapter 2 addressed one of the main problems in hyperspectral pan-sharpening, i.e., the limited number of bands that can be pan-sharpened. Accordingly, an adaptive band selection method was proposed, which involved selecting those bands that are highly correlated with the panchromatic band, instead of selecting the bands that fall inside the panchromatic spectral range. As a lateral investigation, the flexibility of different pan-sharpening methods in accepting bands from different spectral ranges was compared.

- Chapter 3 dealt with the degraded spatial quality problem in the outcomes of the existing RGBW demosaicking methods. It was assumed that the reason for such degradation is that the panchromatic and colour pixels are not effectively involved in each other's demosaicking. Accordingly, a procedure based on a collaborative interpolation between panchromatic and colour pixels was proposed. The performance of the proposed procedure was tested on real RGBW images, beyond the evaluation routine in the literature that uses only simulated images.
- Chapter 4 focused mostly on an unaddressed issue in RGBW demosaicking, i.e., denoising. It presented a deep CNN for RGBW joint demosaicking and denoising. However, when using deep CNNs, another challenge emerged, and that was the lack of ground-truth panchromatic images to be used as the training dataset alongside the RGB channels. To solve this problem, it was proposed to establish a linear transformation to convert the widely available RGB datasets into panchromatic images. The performance of the proposed procedure was again evaluated by using both real and simulated RGBW images.

5.2 Contributions

Improving the spatial quality of the images significantly extends the scope of applications where an image can be used. Overall, this dissertation contributes to further improving the image spatial quality in two areas that had issues in this regard: hyperspectral pan-sharpening in RS, and RGBW demosaicking in terrestrial applications.

For hyperspectral images,

- previously, the spatial quality enhancement of the hyperspectral images through pan-sharpening was restricted to a very small number of bands. It is now feasible to pan-sharpen more hyperspectral bands, even those beyond the panchromatic spectral range, using the proposed adaptive band selection (chapter 2).
- The adaptive band selection did not introduce significant spectral distortions, nor did it disturb the spectral signatures which are used for identifying different land cover types in afterwards applications such as classification.
- This research helps one to make a better decision for choosing a pan-sharpening method that can provide more hyperspectral bands with improved spatial quality without causing severe spectral distortion.
- With many more spatially enhanced bands, a new dataset with higher spatial specifications compared to the original hyperspectral images can be generated. It is expected that the new dataset, would open up new fields of applications for hyperspectral images.

For RGBW demosaicked images,

- the proposed procedure in Chapter 3 significantly improves the spatial quality of the colour images obtained from RGBW CFAs, which was previously highly degraded.

- The visual perception of the images is much enhanced by minimizing zipper effects, colour artifacts and noise, and retrieving finer spatial details, while previously, the images had many quality issues in these regards.
- The research in chapter 4 jointly brings denoising into RGBW demosaicking, allowing for the retrieval of high spatial quality images even in highly noisy conditions. Previously, no denoising was involved in the RGBW demosaicking.
- The visual perception of images captured in noisy situations is highly improved. Previously, the images were heavily contaminated by noise in noisy settings, making the RGBW sensors even less useful.
- Finally, with the improvements made to the image quality, this research increases the potential of bringing RGBW sensors into real-world applications in a variety of noisy situations.

5.3 Suggestions for Future Works

For hyperspectral images,

- the effect of the proposed adaptive band selection can be further investigated for applications such as classification.
- Future studies can investigate the usefulness of the newly created dataset (with many more spatially enhanced bands) to meet new needs that require both high spectral and high spatial qualities at the same time.

For RGBW demosaicked images,

- integrating deep CNN models for RGB demosaicking into the demosaicking-and-fusion strategy in chapter 3 can be considered as another approach to follow. Since the deep networks for demosaicking RGB CFAs have been widely investigated and trained using massive available RGB datasets, they can be used for demosaicking the colour image, while the panchromatic image can be achieved in the same manner as chapter 3.
- The joint demosaicking and denoising model in chapter 4 can be further improved by making the network robust against more complex types of noise that happen in real-world imaging. This can be done by considering a wider noise level range and other noise models such as Poisson.
- Improving the panchromatic simulation required for training the proposed network can be another research topic. One idea could be to design a shallow network to model the relationship between panchromatic and colour images using the images captured by various colour and panchromatic sensors.
- Employing other deep learning networks such as Generative Adversarial Network (GAN) for a joint demosaicking and denoising of RGBW images and comparing with the proposed network can be studied in future.
- Comparing RGB and RGBW CFAs through the performance comparison of the related demosaicking approaches at varying noise levels can be suggested as another direction to be pursued.
- Future studies may compare the demosaicking-and-fusion strategy to the joint demosaicking and denoising approach proposed in chapters 3 and 4, respectively, using more datasets captured under different noisy conditions.

On real RGBW images, it is expected that the first approach would retain higher spatial details in noise free and moderately noisy images, while the second approach is likely to provide higher SNR in moderately to very noisy settings.

References

Aiazzi, B, Alparone, L., Baronti, S., Garzelli, A., & Selva, M. (2006). *MTF-tailored Multiscale Fusion of High-resolution MS and Pan Imagery*. *72*(5), 591–596.

Aiazzi, Bruno, Baronti, S., & Selva, M. (2007). Improving component substitution pansharpening through multivariate regression of MS+ Pan data. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(10), 3230–3239. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4305344

Alparone, L., Aiazzi, B., Baronti, S., Garzelli, A., Nencini, F., & Selva, M. (2008). Multispectral and panchromatic data fusion assessment without reference. *Photogrammetric Engineering and Remote Sensing*, *74*(2), 193–200. http://beta.asprs.org/a/publications/pers/2008journal/february/2008_feb_193-200.pdf

AND9180/D-TRUESENSE Sparse Color Filter Pattern. (2014). *ON Semiconductor Components Industries*. 1–10.

Bayer, B. E. (1976). Color Imaging Array. In *U. S. patent 3971 065*. http://www.google.de/patents/US3971065

Bobcat Intelligent camera series IGV-B2320. (2013). 5–6.

Borman, S., & Stevenson, R. L. (n.d.). *Super-Resolution from Image Sequences - A Review*.

Cetin, M., & Musaoglu, N. (2009). Merging hyperspectral and panchromatic image data: qualitative and quantitative analysis. *International Journal of Remote Sensing*, *30*(7), 1779–1804. https://doi.org/10.1080/01431160802639525

Chavez, P. S., & Kwarteng, A. Y. (1989). Extracting Spectral Contrast in Landsat

Thematic Mapper Image Data Using Selective Principal Component Analysis. *Photogramm. Eng. Remote Sens. No. 3, Pp.*, *55*(3), 339–348.

Chen, Y., Liu, L., Phonevilay, V., Gu, K., Xia, R., Xie, J., Zhang, Q., & Yang, K. (2021). Image super-resolution reconstruction based on feature map attention mechanism. *Applied Intelligence*, *51*(7), 4367–4380. https://doi.org/10.1007/s10489-020-02116-1

Chen, Z., Pu, H., Wang, B., & Jiang, G. M. (2014). Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods. *IEEE Geoscience and Remote Sensing Letters*, *11*(8), 1418–1422. https://doi.org/10.1109/LGRS.2013.2294476

Choi, J., Yu, K., & Kim, Y. (2011). A New Adaptive Component-Substitution-Based Satellite Image Fusion by Using Partial Replacement. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(1), 295–309. https://doi.org/10.1109/TGRS.2010.2051674

Chung, K. L., Chan, T. H., & Chen, S. N. (2020). Effective three-stage demosaicking method for rgbw cfa images using the iterative error-compensation based approach. *Sensors (Switzerland)*, *20*(14), 1–12. https://doi.org/10.3390/s20143908

Clevert, D. A., Unterthiner, T., & Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (ELUs). *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 1–14.

Condat, L. (2014). A generic proximal algorithm for convex optimization - Application to total variation minimization. *IEEE Signal Processing Letters*, *21*(8), 985–989. https://doi.org/10.1109/LSP.2014.2322123

Condat, L. (2009). A GENERIC VARIATIONAL APPROACH FOR DEMOSAICKING FROM AN ARBITRARY COLOR FILTER ARRAY. *16th IEEE International Conference on Image Processing (ICIP)*, *3*, 1–4.

Condat, L. (2012). Joint Demosaicing and Denoising by Total Variation Minimization. *ICIP*, 1–4. https://doi.org/10.1007/s13398-014-0173-7.2

Din, S., Paul, A., & Ahmad, A. (2020). Lightweight deep dense Demosaicking and Denoising using convolutional neural networks. *Multimedia Tools and Applications*, *79*(45–46), 34385–34405. https://doi.org/10.1007/s11042-020-08908-4

Dong, N., Maggioni, M., Yang, Y., Pérez-Pellitero, E., Leonardis, A., & McDonagh, S. (2021). *Residual Contrastive Learning for Joint Demosaicking and Denoising*. http://arxiv.org/abs/2106.10070

Dong, W., Yuan, M., Li, X., & Shi, G. (2018). Joint demosaicing and denoising with perceptual optimization on a generative adversarial network. *ArXiv*, 1–12.

Dozat, T. (2016). Incorporating Nesterov Momentum into Adam. *ICLR Workshop*, *1*, 2013–2016.

Elgendy, O. A., Gnanasambandam, A., Chan, S. H., & Ma, J. (2021). Low-Light Demosaicking and Denoising for Small Pixels Using Learned Frequency Selection. *IEEE Transactions on Computational Imaging*, *7*, 137–150. https://doi.org/10.1109/TCI.2021.3052694

Farrell, J., Xiao, F., & Kavusi, S. (2006). Resolution and light sensitivity tradeoff with pixel size. *Proc. SPIE, Digital Photography II;*, *6069*. https://doi.org/https://doi.org/10.1117/12.646805

Fathollahi, F., & Zhang, Y. (2020). Adaptive band selection for pan-sharpening of hyperspectral images. *International Journal of Remote Sensing*, *41*(10), 3924–3947. https://doi.org/10.1080/01431161.2019.1711240

Fletcher, E. (2018). *Fully-Arbitrary Learned Color Filter Mosaics*. https://evanfletcher42.com/2018/09/23/arbitrary-learned-mosaic/

Gharbi, M., Chaurasia, G., Paris, S., & Durand, F. (2016). Deep joint demosaicking and denoising. *ACM Transactions on Graphics*, *35*(6), 1–12. https://doi.org/10.1145/2980179.2982399

Griffin, M. K., May-Hsu, S., Burke, H. K., Orloff, S. M., & Upham, C. a. (2005). Examples of EO-1 Hyperion data analysis. *Lincoln Laboratory Journal*, *15*(2), 271–298.

Hamilton, J. F., & Compton, J. T. (2007). *PROCESSING COLOR AND PANCHROMATIC PIXELS*. https://doi.org/US 2010/0311130 Al

He, K., Sun, J., & Tang, X. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(6), 1397–1409. https://doi.org/10.1109/TPAMI.2012.213

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Jabari, S., Fathollahi, F., Roshan, A., & Zhang, Y. (2017). Improving UAV imaging quality by optical sensor fusion: an initial study. *International Journal of Remote Sensing*, *38*(17), 4931–4953. https://doi.org/10.1080/01431161.2017.1328145

Jin, Q., Facciolo, G., & Morel, J. M. (2020). A review of an old dilemma: Demosaicking first, or denoising first? *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2020-June*, 2169–2179. https://doi.org/10.1109/CVPRW50498.2020.00265

Kang, M., & Jung, M. (2020). Low-dimensional manifold model for demosaicking from a RGBW color filter array. *Signal, Image and Video Processing*, *14*(1), 143–150. https://doi.org/10.1007/s11760-019-01535-z

Khadidos, A. O., Khadidos, A. O., Khan, F. Q., Tsaramirsis, G., & Ahmad, A. (2021). Bayer image demosaicking and denoising based on specialized networks using deep learning. *Multimedia Systems*, *27*(4), 807–819. https://doi.org/10.1007/s00530-020-00707-z

Khiabani, S. A. (2015). *Terrestrial Implementation of UNB Super Camera And Improvements To UNB-Pansharp* (Issue May) [UNIVERSITY OF NEW BRUNSWICK]. https://unbscholar.lib.unb.ca/islandora/object/unbscholar%3A7107/datastream/PDF/view

Kijima, T., Nakamura, H., Compton, J. T., Hamilton, J. F., & DeWeese, T. E. (2006). *IMAGE SENSOR WITH IMPROVED LIGHT SENSITIVITY*.

Kiku, D., Monno, Y., Tanaka, M., & Okutomi, M. (2014). Minimized-Laplacian Residual Interpolation for Color Image Demosaicking. *Proc. OfSPIE, 9023,* 90230L–1–8. https://doi.org/10.1109/TIP.2016.2518082

Kiku, D., Monno, Y., Tanaka, M., & Okutomi, M. (2016). Beyond Color Difference: Residual Interpolation for Color Image Demosaicking. *IEEE Transactions on Image Processing*, *25*(3), 1288–1300. https://doi.org/10.1109/TIP.2016.2518082

Kiku, D., Monno, Y., Tanaka, M., & Okutomi, M. (2013). RESIDUAL INTERPOLATION FOR COLOR IMAGE DEMOSAICKING. *IEEE Int. Conf. on Image Processing (ICIP)*, *2*, 2304–2308.

Kim, H., Lee, S., & Kang, M. G. (2020). Demosaicing of RGBW color filter array based on rank minimization with colorization constraint. *Sensors (Switzerland)*, *20*(16), 1–18. https://doi.org/10.3390/s20164458

Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 1646–1654. https://doi.org/10.1109/CVPR.2016.182

Kimmel, R. (1999). Demosaicing: Image reconstruction from color CCD samples. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, *8*(9), 1221–1228. https://doi.org/10.1007/BFb0055693

Kokkinos, F., & Lefkimmiatis, S. (2019). Iterative joint image demosaicking and denoising using a residual denoising network. *IEEE Transactions on Image Processing*, *28*(8), 4177–4188. https://doi.org/10.1109/TIP.2019.2905991

Kwan, C., & Chou, B. (2019). Further improvement of debayering performance of RGBW color filter arrays using deep learning and pansharpening techniques. *Journal of Imaging*, *5*(8). https://doi.org/10.3390/jimaging5080068

Kwan, C., Chou, B., Kwan, L.-Y. M., & Budavari, B. (2017). Debayering RGBW Color Filter Arrays : A Pansharpening Approach. *IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference*.

Kwan, C., & Larkin, J. (2019). Demosaicing of bayer and cfa 2.0 patterns for low lighting images. In *Electronics (Switzerland)* (Vol. 8, Issue 12). https://doi.org/10.3390/electronics8121444

Laben, C., & Brower, B. (2000). Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. *Tech. Rep. US Patent*, *6,011,875*. http://www.google.com/patents?hl=en&lr=&vid=USPAT6011875&id=5v0AAAAAEBAJ&oi=fnd&dq=Process+for+enhancing+the+spatial+resolution+of+multispectral+imagery+using+pan-sharpening&printsec=abstract

LeCun, Y. A., Bottou, L., Orr, G. B., & Müller, K.-R. (1998). Efficient BackProp. In *Neural Networks: Tricks of the Trade* (2nd ed.). Springer.

Liang, S., & Wang, J. (Eds.). (2020). Chapter 1 - A systematic view of remote sensing,. In *Advanced Remote Sensing: Terrestrial Information Extraction and Applications* (Second Edi, pp. 1–57). Academic Press. https://doi.org/https://doi.org/10.1016/C2017-0-03489-4

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision -- ECCV 2014* (pp. 740–755). Springer International Publishing.

Liu, J. G. (2000). Smoothing Filter-based Intensity Modulation: a spectral preserve image fusion technique for improving spatial details. *Int. j. Remote Sensing*, *21*(18), 3461–3472.

Loncan, L., Almeida, L. B., Bioucas-Dias, J. M., Briottet, X., Chanussot, J., Dobigeon, N., Fabre, S., Liao, W., Licciardi, G. A., Simões, M., Tourneret, J.-Y., Veganzones, M. A., Vivone, G., Wei, Q., & Yokoya, N. (2015). Hyperspectral pansharpening: a review. *IEEE Geoscience and Remote Sensing Magazine*, *3*(3), 27–46. https://doi.org/10.1109/MGRS.2015.2440094

Losson, O., Macaire, L., & Yang, Y. (2010). Comparison of color demosaicing methods. *Advances in Imaging and Electron Physics*, *162*(C), 173–265. https://doi.org/10.1016/S1076-5670(10)62005-8

Ma, K., Duanmu, Z., Wu, Q., Wang, Z., Yong, H., Li, H., & Zhang, L. (2017). Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, *26*(2), 1004–1016. https://doi.org/10.1109/TIP.2016.2631888

Malvar, H. S., He, L. H. L., & Cutler, R. (2004). High-quality linear interpolation for demosaicing of Bayer-patterned color images. *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, *3*, 5–8. https://doi.org/10.1109/ICASSP.2004.1326587

Meng, X., Shen, H., Li, H., Zhang, L., & Fu, R. (2019). Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges. *Information Fusion*, *46*(June 2018), 102–113. https://doi.org/10.1016/j.inffus.2018.05.006

Menon, D., & Calvagno, G. (2011). Color image demosaicking: An overview. *Signal Processing: Image Communication*, *26*(8–9), 518–533. https://doi.org/10.1016/j.image.2011.04.003

Moeller, M., Wittman, T., & Bertozzi, A. L. (2009). A variational approach to hyperspectral image fusion. *In SPIE Defense, Security, and Sensing International Society for Optics and Photonics*, 73341E-73341E – 10.

https://doi.org/10.1117/12.818243

Monno, Y., Kiku, D., Tanaka, M., & Okutomi, M. (2017). Adaptive Residual Interpolation for Color and Multispectral Image Demosaicking. *Sensors*, *17*(12). https://doi.org/10.3390/s17122787

Palubinskas, G., & Reinartz, P. (2011). Multi-resolution, multi-sensor image fusion: general fusion framework. *Joint Urban Remote Sensing Event (JURSE)*, 313–316. https://doi.org/10.1109/JURSE.2011.5764782

Pan, Z., Li, B., Cheng, H., & Bao, Y. (2020). Joint Demosaicking and Denoising for CFA and MSFA Images Using a Mosaic-Adaptive Dense Residual Network. In A. Bartoli & A. Fusiello (Eds.), *Computer Vision -- ECCV 2020 Workshops* (pp. 647–664). Springer International Publishing.

Park, B., & Jeong, J. (2019). Color filter array demosaicking using densely connected residual network. *IEEE Access*, *7*, 128076–128085. https://doi.org/10.1109/ACCESS.2019.2939578

Picone, D., Restaino, R., Vivone, G., Addesso, P., Dalla Mura, M., & Chanussot, J. (2017). Band Assignment Approaches for Hyperspectral Sharpening. *IEEE Geoscience and Remote Sensing Letters*, *14*(5), 739–743. https://doi.org/10.1109/LGRS.2017.2677087

Rafinazari, M. (2017). *Denoising and Demosaicking of Color Images*. University of Ottawa.

Rafinazari, M., & Dubois, E. (2016). Demosaicking algorithms for RGBW color filter arrays. *IS&T International Symposium on Electronic Imaging Color Imaging XXI: Displaying, Processing, Hardcopy, and Applications*, *c*, 1–6.

Simoes, M., Bioucas-Dias, J., Almeida, L. B., & Chanussot, J. (2015). A convex formulation for hyperspectral image superresolution via subspace-based regularization. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(6), 3373–3388. https://doi.org/10.1109/TGRS.2014.2375320

Singh, M., & Singh, T. (2012). LINEAR UNIVERSAL DEMOSAICKING OF REGULAR PATTERN COLOR FILTER ARRAYS. *IEEE International Conference on Acoustics, Speech and Signal (ICASSP)*, *5*, 1277–1280. https://doi.org/10.1109/ICASSP.2012.6288122

Syu, N. S., Chen, Y. S., & Chuang, Y. Y. (2018). Learning Deep Convolutional Networks for Demosaicing. *ArXiv*, 1–14.

Tan, D. S., Chen, W. Y., & Hua, K. L. (2018). DeepDemosaicking: Adaptive Image Demosaicking via Multiple Deep Fully Convolutional Networks. *IEEE Transactions on Image Processing*, *27*(5), 2408–2419. https://doi.org/10.1109/TIP.2018.2803341

Tan, H., Xiao, H., Liu, Y., & Zhang, M. (2022). Two-Stage CNN Model for Joint Demosaicing and Denoising of Burst Bayer Images. *Computational Intelligence and Neuroscience*, *2022*.

Tu, T. M., Su, S. C., Shyu, H. C., & Huang, P. S. (2001). A new look at IHS-like image fusion methods. *Information Fusion*, *2*(3), 177–186. http://www.sciencedirect.com/science/article/pii/S1566253501000367

Vivone, G., Alparone, L., Chanussot, J., Mura, M. D., Garzelli, A., Licciardi, G. A., Restaino, R., & Wald, L. (2015). A Critical Comparison Among Pansharpening Algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(5), 2565–2586.

Vivone, G., Restaino, R., Licciardi, G., Dalla Mura, M., & Chanussot, J. (2014). MULTIRESOLUTION ANALYSIS AND COMPONENT SUBSTITUTION TECHNIQUES FOR HYPERSPECTRAL PANSHARPENING. *IEEE IGARSS*, 2649–2652. https://doi.org/10.1021/bi00827a011

Vivone, G., Restaino, R., Mura, M. D., Licciardi, G., & Chanussot, J. (2014). Contrast and error-based fusion schemes for multispectral image pansharpening. *IEEE Geoscience and Remote Sensing Letters*, *11*(5), 930–934. https://doi.org/10.1109/LGRS.2013.2281996

Wang, J., Zhang, C., & Hao, P. (2011). NEW COLOR FILTER ARRAYS OF HIGH LIGHT SENSITIVITY AND HIGH DEMOSAICKING PERFORMANCE. *Ieee International Conference On Image Processing*, *0*, 3214–3217. https://doi.org/10.1109/ICIP.2011.6116336

Wang, L., Wang, Y., Dong, X., Xu, Q., Yang, J., An, W., & Guo, Y. (2021). Unsupervised Degradation Representation Learning for Blind Super-Resolution. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 10576–10585. https://doi.org/10.1109/CVPR46437.2021.01044

Wei, Q. (2015). *Bayesian Fusion of Multi-band Images: A Powerful Tool for Super-resolution*. UNIVERSIT´E DE TOULOUSE.

Wei, Q., Bioucas-Dias, J., Dobigeon, N., & Tourneret, J.-Y. (2016). Hyperspectral and Multispectral Image Fusion based on a Sparse Representation. *IEEE J. Sel. Topics Signal Process*, 72–76. https://doi.org/10.1109/TALE.2015.7386019

Wu, F., Huang, T., Dong, W., Shi, G., Zheng, Z., & Li, X. (2021). Toward blind joint demosaicing and denoising of raw color filter array data. *Neurocomputing*, *453*, 369–382. https://doi.org/10.1016/j.neucom.2020.09.090

Xu, Q., Zhang, Y., & Li, B. (2014). Recent advances in pansharpening and key problems in applications. *International Journal of Image and Data Fusion*, *5*(3), 37–41. https://doi.org/10.1080/19479832.2014.889227

Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Transactions on Multimedia*, *21*(12), 3106–3121. https://doi.org/10.1109/TMM.2019.2919431

Yokoya, N., Yairi, T., & Iwasaki, A. (2012). Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(2), 528–537. https://doi.org/10.1109/TGRS.2011.2161320

Zhang, C., Li, Y., Wang, J., & Hao, P. (2016). Universal demosaicking of color filter arrays. *IEEE Transactions on Image Processing*, *25*(11), 5173–5186. https://doi.org/10.1109/TIP.2016.2601266

Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, *26*(7), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206

Zhang, L., Wu, X., Budaes, A., & LI, X. (2011). Color demosaicking by local directional interpolation and nonlocal adaptive thresholding. *Journal of Electronic Imaging*, *20*(2), 023016. https://doi.org/10.1117/1.3600632

Zhang, Y. (2008). Pan-sharpening for improved information extraction. *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS*

Congress Book, 185–203. https://doi.org/10.1201/9780203888445.ch14

Zhang, Y., & Mishra, R. K. (2014). From UNB PanSharp to Fuze Go - the success behind the pan-sharpening algorithm. *International Journal of Image and Data Fusion*, *5*(1), 39–53. https://doi.org/10.1080/19479832.2013.848475

Curriculum Vitae

Candidate's full name: Fatemeh Fathollahi Kalanpa

Universities attended (with dates and degrees obtained):

- 2013: M.Sc., Remote Sensing, Department of Geomatics Engineering, University of Tehran, Iran
- 2009: B.Sc., Geomatics Engineering, Department of Geomatics Engineering, University of Tehran, Iran

Publications:

Fathollahi, Rezaee, M F., & Zhang, Y. (2021) Deep Joint Demosaicking and Denoising for RGBW Colour Filter Arrays. *IEEE Transaction on computational Imaging.*, Submitted.

Fathollahi, F., Jabari, S., Zhang, Y., & Rezaee, M. (2021) RGBW demosaicking using a collaborative interpolation between panchromatic and colour pixels. *Signal Processing: Image comunication,*Under review.

Fathollahi, F. & Zhang, Y. (2020). Adaptive band selection for pan-sharpening of hyperspectral images. *International Journal of Remote Sensing*, *41*(10)

Jabari, S., Rezaee, M., Fathollahi, F., & Zhang, Y. (2019). Multispectral change detection using multivariate Kullback-Leibler distance. *ISPRS Journal of Photogrammetry and Remote Sensing*, *147*.

DadrasJavan, F., Samadzadegan, F., & Fathollahi, F. (2018). Spectral and Spatial Quality Assessment of IHS and Wavelet Based Pan-sharpening Techniques for High Resolution Satellite Imagery. *Advances in Image and Video Processing*, *6*(2).

Jabari, S., Fathollahi, F., & Zhang, Y. (2017). Application of sensor fusion to improve UAV image classification. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42* (2W6), 153–156.

Jabari, S., Fathollahi, F., Roshan, A., & Zhang, Y. (2017). Improving UAV imaging quality by optical sensor fusion: An initial study. *International Journal of Remote Sensing*, *38*(17), 4931–4953.

Zhang, Y., Roshan, A., Jabari, S., Khiabani, S. A., Fathollahi, F., & Mishra, R. K. (2016). Understanding the quality of pansharpening – A lab study. *Photogrammetric Engineering and Remote Sensing*, *82*(10).

Adham Khiabani, S. A., Zhang, Y., & Fathollahi, F. (2014). Influence of image compression on the quality of UNB pan-sharpened imagery: A case study with security video image frames. *Proceedings of SPIE - The International Society for Optical Engineering*, 9100.

Adham Khiabani, S., Zhang, Y., & Fathollahi, F. (2014). Image quality assessment of 2-chip color camera in comparison with 1-chip color and 3-chip color cameras in various lighting conditions: initial results. Image Sensing Technologies: Materials, Devices, Systems, and Applications, 9100.

Conference Presentations:

Fathollahi, F. Rezaee, M., & Zhang, Y. (2019). RGBW joint demosaicking and denoising using a CNN. *Presented at the Int. Conf. PIERS, Xiamen, China.*

Fathollahi, F. & Zhang, Y. (2017). Improving Spectral Quality Assessment of the PanSharpened Images using IHS-based SAM and SCM. *Presented at the Int. Conf. ASPRS IGTF, Baltimore, USA.*

Fathollahi, F., Zhang, Y., Jabari, S., & Rezaee, M (2017). Improving the Colour Distortion of the Hyperspectral Pan-Sharpening by Incorporating More Spectral Bands. *Presented at the Int. Conf. ASPRS IGTF, Baltimore, USA.*

Samadzadegan, F., Fathollahi, F., & Rezaee, M. (2012). Comparing different IHS-based pan-sharpening techniques for worldview-2 high resolution satellite imagery. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2296–2299.