Stitching from Spectral Filter Array Video Sequences

Abdelhamid N. Fsian[1(✉)], Jean-Baptiste Thomas[1,2], Jon Y. Hardeberg[2], and Pierre Gouton[1]

[1] ImVia, Université de Bourgogne, Dijon, France
abdelhamidfsian@gmail.com
[2] Colourlab, Department of Computer Science, NTNU, Gjøvik, Norway

Abstract. Hyperspectral imaging offers high spectral and spatial resolution, but its high costs and time-consuming nature make it difficult to use. Spectral Filter Array (SFA) imaging presents an alternative, offering high spectral resolution, user-friendliness, and affordability, but at the cost of limited spatial resolution. This paper presents an approach to address this trade-off, starting with raw overlapping frames from spectral videos, followed by a demosaicking network process before tackling the stitching problem. Our experiments on various spectral videos, supported by image quality metrics and qualitative demonstrations, indicate that this approach effectively enhances the spatial resolution of spectral images while reducing artifacts. The integration of the demosaicking and the stitching provides a robust solution for spectral video applications, paving the way for further advancements in panoramic spectral image stitching.

Keywords: Spectral Imaging · Stitching · Demosaicking · Super Resolution

1 Introduction

Spectral Imaging holds immense potential for applications such as food safety inspection [26], land cover categorization [16] and object tracking [12], offering a nuanced and comprehensive perspective that extends beyond the capabilities of traditional imaging techniques. In the realm of cultural heritage, the utilization of spectral videos can prove instrumental in capturing intricate details of artifacts, manuscripts, and historical sites [24]. But due to hardware limitations, the spatial resolution is limited, compared to Color Filter Array (CFA) [3] imaging. Moreover, demosaicking which refers to the process of creating a fully-defined spectral image free of spatial and spectral distortions, is a crucial stage for Spectral Filter Array (SFA) camera [13]. Through the seamless integration of demosaicking and stitching algorithms, spectral videos can be transformed into large panoramas by combining multiple spectral images with overlapping areas. Demosaicking, in particular, plays a pivotal role in reconstructing a fully-defined

R. Schettini et al. (Eds.): CCIW 2024, LNCS 15193, pp. 132–146, 2025.
https://doi.org/10.1007/978-3-031-72845-7_10

image from a mosaic image. This process enhances the visual clarity and accuracy of the final image, enabling the preservation of spectral informations [18].

In the context of remote sensing, spectral video holds the key to unlocking a wealth of information about the Earth's surface [29]. By capturing data across multiple spectral bands, these videos provide a richer and more comprehensive understanding of the environment. The combination of demosaicking and stitching techniques in remote sensing applications facilitates the creation of high-resolution, spectral panoramas. This panoramic view can be instrumental in monitoring land cover changes, assessing vegetation health, and identifying environmental anomalies.

By stitching together multiple images of the same location taken from various angles, a panorama image is created through the process of image stitching. In general, when dealing with an issue involving just two images, one of the images is used as the reference, while the other is used to estimate a 3×3 homography, also known as a camera matrix [11]. However, the more images that are included, the more complex this problem becomes [22]. Another pivotal consideration in the construction of panoramas involves not only the spatial arrangement of images but also the harmonization of visual attributes across individual frames. It is imperative to adjust the appearance of each image meticulously to eliminate discernible boundaries in the resultant panorama. Even when images are captured simultaneously, variations in exposure, contrast, and the introduction of effects such as vignetting may necessitate careful calibration to achieve seamless integration [31]. Given the context of working with spectral images, the meticulous harmonization of visual attributes to achieve seamless transitions between images can result in a reduction of the spectral fidelity. This unintended consequence is undesirable, emphasizing the need for caution when undertaking the harmonization of visual attributes.

This paper aims to contribute to the field of spectral imaging by exploring the potential of spectral stitching to enhance the quality and spatial resolution of spectral images. By increasing the number of spectral images used in the stitching process, it is possible to achieve a more detailed and accurate representation of the target object or scene. This approach has the potential to be particularly useful in applications such as remote sensing, cultural heritage preservation, where the ability to capture highly detailed spectral information is crucial. This paper presents our findings on how the number of spectral images impacts the quality and spatial resolution of stitched images.

This paper is structured as follows: Sect. 2 offers an overview of existing literature on stitching algorithms and spectral demosaicking. In Sect. 3, we delve into our proposed methodology, providing a detailed explanation of our approach. Section 4 outlines our experimental setup and presents our analysis of the obtained results. Finally, we conclude with a brief summary of our findings and discuss avenues for future research.

2 Related Work

SFA camera emerged as a good option to captures well-aligned spectral images using a one-shot acquisition [13]. An SFA is made by arranging several bandpass optical filters in a pattern that repeats spatially, covering the monochromatic imaging sensor, inspired from the Bayer CFA [3]. Since all spectrum bands in the SFA are spatially subsampled, only sparse data from a single band is preserved in each individual pixel of the collected original spectral image. Consequently, the missing spectral values must be approximated from the gathered sparse spatial data in order to obtain fully-defined (demosaicked) spectral images (see Fig. 1). This process is referred to as SFA demosaicking [5], provides a critical component for reconstructing spectral images for SFA based camera [21].

Several handcrafted strategies have been investigated in previous research on SFA image demosaicking [25,27]. The weighted bilinear (WB) interpolation approach was presented by Brauers and Aach [5], extending CFA demosaicking to SFA. Chini et al. [7] first introduced the idea of the pseudo-panchromatic image (PPI), which is the average image of all spectral channels. The PPI difference (PPID) approach was developed by Mihoubi et al. [17,18] to improve SFA demosaicking. It sharpens the PPI by taking the channel residual structure and spatial-spectral correlation of SFA images into account. Convolutional neural networks (CNNs) have demonstrated potential in demosaicking tasks in recent years. Using mosaic convolution-attention networks, Feng et al. [10] developed a deep CNN that focuses on first feature extraction from unprocessed mosaic images. Nevertheless, checkerboard distortion results from their mosaic convolution module ignoring for spatial location variations. The demosaicking architecture was updated by Pan et al. [15] by utilizing a conventional two-pass residual interpolation approach and adding PPIs computed by a CNN. Despite its effectiveness, this method might have halo effects when guided filters are used. By introducing a Residual Network and improving PPI generation, Zhao et al. [30] built upon the work of Pan et al. [15]. They achieved state-of-the-art metrics and enhanced demosaicking outcomes by including edge-related data and using adaptive spatial and spectral correction inside the network.

Furthermore, image stitching has been an active area of research for many years, with methods continuously becoming more robust and yielding better results. Early methods relied on user-provided information to produce visually appealing panoramas. A major paper named AutoStitch by Matthew Brown and David G. Lowe [6], along with their proprietary software tool that could allow anyone to produce great panoramas simply by uploading their images in any order. Their algorithm was based on the features detector and descriptor SIFT [8], proposed by Lowe in 2004, and the RANSAC algorithm [4]. Since then, multiple other methods have emerged, each one proposing better results in edge cases by computing the stitching in a slightly different manner. Among the most popular recent methods, we can mention the APAP [28] and the AANAP [14] algorithms. We propose in this paper to implement a python version of the AutoStitch [6] algorithm specifically for spectral imaging. Since it is a well-established algorithm in the stitching community.

3 Methodology

Starting with the initial phase, our approach integrates a demosaicking network to obtain fully defined spectral images from mosaic images. This foundational step ensures refined inputs for the subsequent image stitching process. Transitioning to image stitching, the absence of ground truth image makes it impossible to apply metrics like PSNR or SSIM, as there is no reference for comparison between the stitched image and a target image. To this end SOTA stitching solution [14,28], often uses blind image quality assessment metrics such as BRISQUE [19] or NIQUE [20].

Our sequential procedure begins with the application of our Deep Spectral Demosaic network to obtain fully-defined (demosaicked) spectral images. Next, we apply the SIFT [9] and RANSAC [4] algorithms to the PPI for feature matching. To address variations in gain, we integrate gain compensation techniques and perform multi-band blending to seamlessly merge the images. To obtain a spectral stitched image, we apply the same computed parameters for the other bands without any new parameter computations, as all the bands have the same field of view.

3.1 Deep Demosaic Net

The proposed deep learning model comprises two integral modules: the Deep PPI module (PPI-net) and the Deep Residual Demosaic module (DRDm). The PPI-net is dedicated to extracting a pseudo-panchromatic image (PPI) from a raw mosaic image, serving as an intermediate representation capturing spatial information from multiple channels. Pre-trained to reconstruct a pseudo-panchromatic image from the raw SFA image, the PPI-net establishes a foundational step (see Fig. 1).

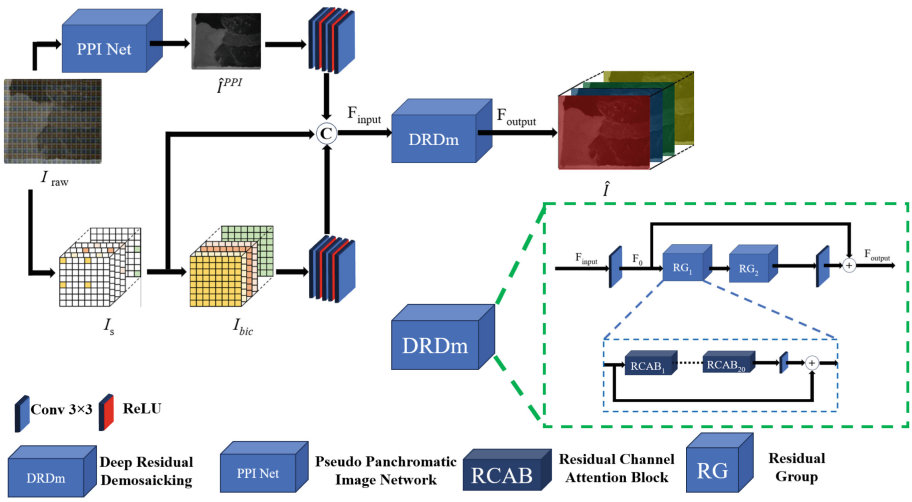


Fig. 1. Deep Spectral Demosaicking network architecture

The proposed deep demosaicking model comprises a two-branch CNN, the first branch consists of a deep pseudo-panchromatic network (PPI-net) in order to estimate the pseudo-panchromatic image from raw image, drawing inspiration from the successful methodology proposed by [23,30]. Additionally, the second branch serves as a preliminary demosaicking outcome achieved through bicubic interpolation. These branches are seamlessly combined into a cohesive module, namely the deep residual demosaicking (DRDm).

PPI-net. The determination of the pseudo-panchromatic image (PPI) at each pixel (p) involves computing the average value across all channels of spectral channels, as outlined in prior research [18].

$$I_p^{PPI} = \frac{1}{N} \sum_{n=1}^{N} I_p^{GT}, \tag{1}$$

where N denote the number of spectral bands.

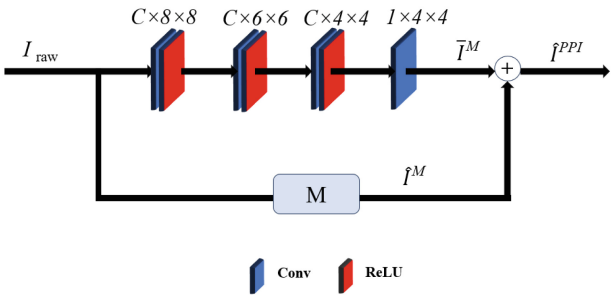


Fig. 2. Illustration of PPI-net module. Where $\oplus$ denotes elementwise sum.

Given the clear positive linear correlation found between the reconstructed spectral image and the preserved high-frequency details in the pseudo-panchromatic image (PPI) [23], we propose employing elementwise summation between $\bar{I}^M$ and $\hat{I}^M$, as depicted in Fig. 2. Here, M represents a filter matrix tailored for a 3×3 SFA pattern [15,18].

$$M = \frac{1}{36} \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix}. \tag{2}$$

Additionally, the PPI-net module's loss function, represented by Eq. 3, is meant to minimize the difference between the estimated PPI produced by the PPI-net and the ground truth PPI determined by Eq. 1.

$$L_{PPI} = \sum_{w=1}^{W} \sum_{h=1}^{H} \left|\left| \hat{I}^{PPI}(h,w) - I^{PPI}(h,w) \right|\right|, \tag{3}$$

where L_{PPI} is using MSE. h and w represents respectively the height and width of the image. Moreover, $\hat{I}^{PPI}$ is the estimated PPI and I^{PPI} is the ground truth PPI.

DRDm Module. Following the calculation of the pre-demosaicked image, using bicubic interpolation denoted as I_{bic} and the estimated PPI denoted as $\hat{I}^{PPI}$, both outputs are passed on into two identical blocks. Following this, they are concatenated and used as the DRDm module's input jointly with the sparsity image. The reconstruction part, shallow feature extraction, residual in residual (RIR) deep feature extraction, and deep feature extraction are the main components of our DRDm module.

Furthermore, the loss function of the DRDm module aim to minimize the difference between the estimated demosaicked spectral image and the ground truth image I^{GT}. Specifically, this loss function is the Mean Squared Error (MSE), which is defined as:

$$L_{DRDm} = \frac{1}{P} \sum_{p=1}^{P} \left|\left| I_p^{GT} - \hat{I}_p \right|\right|_2^2, \tag{4}$$

here, P refers to the total count of pixels, where p denotes the index of the pixel as indicated in Eq. 4.

3.2 Training Details

Our demosaicking model was trained using the ARAD-1K dataset [2]. This framework is implemented based on PyTorch, and trained using a NVIDIA RTX 2080 GPU equipped with 20 GB of VRAM and CUDA 11.8 for a total of 2000 epochs. We employed the Adam optimizer to update model parameters, specifying $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Data augmentation techniques were applied to enhance the robustness of the training process, encompassing random cropping of spectral images from ARAD-1K datasets to 128×128 patches, as well as rotations by 90°, 180°, and 270°. Moreover, the batch size was set to 16. To facilitate convergence, we initialized the learning rate to 10^{-4} and adopted a halving strategy every 250 epochs. Moreover, All testing experiments are implemented using the same machine: Intel Core i9-11900k CPU 2.40 GHz, NVIDIA GPU RTX 3070 and 16 Gb of RAM.

3.3 Image Stitching

Feature Matching. The first step is to characterize each image by extracting local features. To do so, we use the SIFT extractor [8], which for each image gives

us the locations of a certain number of keypoints, along with a descriptor for each of these keypoints. The main reason for using SIFT instead of other methods such as the Harris Corner Detector [9] is that SIFT features are invariant under rotation and scale changes, which means that our model can handle images with varying orientation and zoom. Please note that, from now on, images will be indexed as i and j to differentiate between the two images we are trying to stitch.

Once the features have been extracted from all images, a match is being calculated for each pair of images. To do so, each feature of one image is matched to its two nearest neighbors in feature space from the other image. Then, the Lowe's ratio test is used to keep only relevant matching features. This test consists in comparing the first nearest neighbors to the second, and keeping only the matches where the distance with the closest neighbor is significantly smaller than the distance with the second closest one (significantly being determined by the value of the used scaling factor). This test performs a soft thresholding that eliminates most of the wrong matches. After this step, we have for each pair of images a list of matches, each one composed of a feature from each images. The next step is to use those matches to identify the images that would overlap each other in the panorama, and to compute homography matrices (Fig. 3).

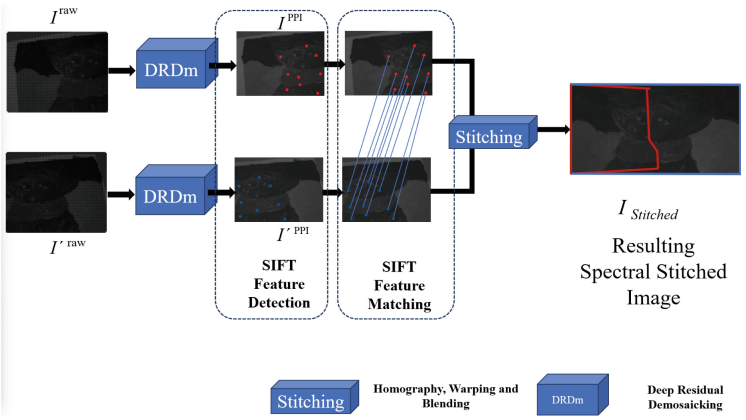


Fig. 3. Overview of our stitching pipeline, where we use only the resulting PPI image to compute stitching parameters. The rest of the spectral bands do not require computing stitching parameters from scratch; instead, the computed stitching parameters from the PPI band can be directly applied to them.

Robust Homography Estimation Using RANSAC. Using a minimum collection of randomly selected correspondences, RANSAC (random sample consensus) [4] is a robust estimation process that estimates images transformation parameters and selects the solution that best corresponds with the data. It works with multiple iterations. At each iteration, a set of four feature matches is randomly selected, and is used to compute the homography H between them using the direct linear transformation method [28]. This operation is repeated a certain number of times, each with a different randomly selected set of four samples.

Then, the homography with the highest number of inliers (i.e. matches whose projections are consistent with H within a tolerance ϵ pixels) is selected. It can be shown that, for a pair of images where half of the matches are corrects, the probability to have not found the correct homography after 500 iterations is approximately 1×10^{-14}, thus making it a pretty robust method.

Probabilistic Model for Verifying Image Matches. At this stage, each pair of images consists of a list of features matches and a homography for one image relatively to the other. However, we do not know yet which pairs of images should indeed overlap in the final panorama, and which pairs actually do not have any area in common. To verify the image matches, we compare the features matches that are consistent with the homography (RANSAC inliers) with the features matches inside the area of overlap between the two images but not consistent (RANSAC outliers). For a given image we denote the total number of features in the area of overlap n_f and the number of inliers n_i. Then, by making small assumptions on the distributions of the feature matches, we can estimate the probability that the image match is correct using these two numbers, and obtain a condition for an image match to be correct:

$$n_i > \alpha + \beta \times n_f. \tag{5}$$

The detailed calculus can be found in the AutoStitch original paper [6]. The parameters α and β could theoretically been adjusted based on the data, but we used the provided values, that is $\alpha = 8$ and $\beta = 0.3$

Using the above condition, we can keep only the correct image matches, and divide the data into connected components (i.e. sets of images that are linked together by matches). This allows us to recognize multiple panoramas in a set of images and reject noisy images that do not match any other images.

3.4 Gain Compensation

We discussed how to calculate each camera's geometric characteristics in earlier sections. Here, we demonstrate how to get the total gain across images, a photometric parameter. Similar configuration applies here, with an error function established across all images. For all overlapping pixels, the error function is the sum of the gain-normalized intensity errors [6]:

$$e = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} N_{ij} \left(\frac{(g_i I_{ij} - g_j I_{ji})^2}{\sigma_N^2} + \frac{(1 - g_i)^2}{\sigma_g^2} \right). \tag{6}$$

where $N_{ij} = |R(i,j)|$ is the number of pixels in image i that overlap in image j , $\bar{I}_{ij}$ is the mean intensity of the image i in the overlap region with image j, defined as

$$\bar{I}_{ij} = \frac{\sum_{u_i \in R(i,j)} I_i(u_i)}{N_{ij}}, \tag{7}$$

where g_i is the gain associated with image I. The parameters σ_N and σ_g denote the standard deviations of the normalized intensity error and gain, respectively. It is noteworthy that an increase in these parameters is expected to result in a reduction of spectral fidelity. This phenomenon is explained by the relatively low standard deviations, specifically $\sigma_N = 1.5$ and $\sigma_g = 0.1$, underscoring the nuanced impact of parameter settings on the fidelity of spectral representations (Fig. 4).

Fig. 4. Left: Stitch Image with multi band blending. Right: Stitch Spectral Image without blending.

3.5 Blending

Multi-band Blending. The idea behind multi-band blending is to blend low-frequency information over a large spatial range, and high-frequency information over a short range. This approach ensures that small details are preserved while still allowing for a smooth transition between images. To achieve this, we first initialize blending weights for each image by finding the set of points for which image i is primarily responsible:

$$W^i_{max} = \begin{cases} 1 & \text{if } W^i = \arg\max W^j \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

in other words, W^i_{max} will be 1 when image i has the maximum weight and 0 where any other image has a higher weight. These maximum-weight maps are progressively blurred to create the blending weights for each band, along with a similar blurring performed on the image itself to build each band. For the first band, a high pass version of the rendered image is formed:

$$I^i_\sigma = I^i * g_\sigma, \tag{9}$$

$$B^i_\sigma = I^i - I^i_\sigma, \tag{10}$$

where $*$ indicates convolution and g_σ is a Gaussian of standard deviation σ. In doing so, the spatial frequencies in the range of $\lambda \subset [0, \sigma]$ are represented by B_σ. The same convolution is used to construct the weight matrix linked to this band

$$W_\sigma^i = W_{max}^i * g_\sigma, \tag{11}$$

here, W_σ^i represent the blend weights Then, other bands are computed iteratively, using a standard deviation $\sigma' = \sqrt{(2k+1)}\sigma$ which is getting larger, with $k > 0$

$$I_{(k+1)\sigma}^i = I_{k\sigma}^i * g_{\sigma'}, \tag{12}$$

$$B_{(k+1)\sigma}^i = I_{k\sigma}^i - I_{(k+1)\sigma}^i, \tag{13}$$

$$W_{(k+1)\sigma}^i = W_{k\sigma}^i * g_{\sigma'}, \tag{14}$$

this way, each band k represents spacial frequencies in the range of wavelengths $\lambda \subset [\sqrt{(2k-1)}\sigma, \sqrt{(2k+1)}\sigma]$. Finally, overlapping images are merged linearly for each band using the associated blend weights.

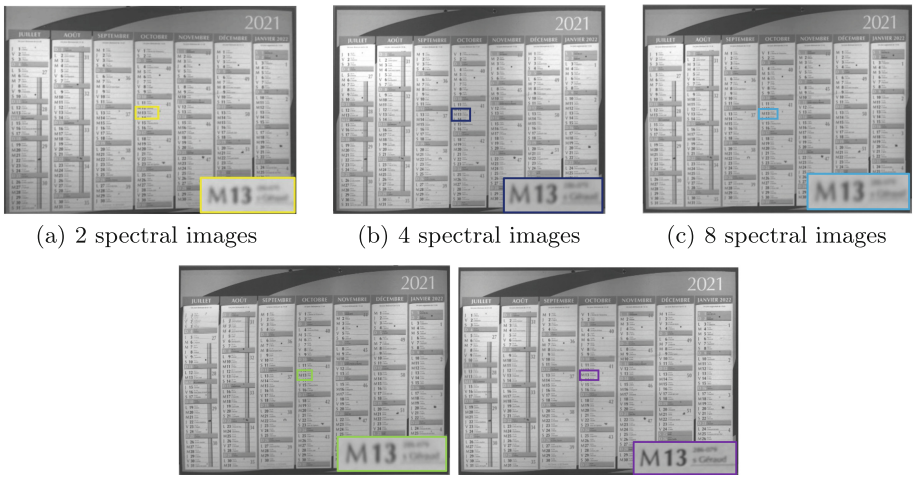
$$I_{k\sigma}^{multi} = \frac{\sum_{i=1}^n B_{k\sigma}^i W_{k\sigma}^i}{\sum_{i=1}^n W_{k\sigma}^i}. \tag{15}$$

More details and illustrations about this method can be found in the AutoStitch paper [6].

4 Experiment

To evaluate the effectiveness of our method, we designed a series of experiments aimed at discerning the impact of both viewing distance and the number of spectral images used on the visual quality of the stitched images. The experiments were conducted under controlled conditions to ensure consistency and reliability of the results. We began by selecting a calibrated object and placing it under uniform D65 illumination. Using a 9 band SFA camera from SILIOS Technologies [1], we captured raw images of the object from a considerable distance manually controlling the focus of the camera. The images were taken in a handheld manner, which introduced small rotations and translations. Initially, only two images were acquired, each covering a portion of the object's field of view. Subsequently, we systematically decreased the distance between the camera and the object, thereby increasing the number of raw images captured for stitching purposes. This approach allowed us to examine the impact of viewing distance on the stitching process and the resulting image quality (see Figs. 5 and 6).

The captured images were processed using our implemented stitching algorithm, which combines multiple images to create a panoramic view of the object. The algorithm utilizes the multi-band blending technique to seamlessly merge



(a) 2 spectral images (b) 4 spectral images (c) 8 spectral images

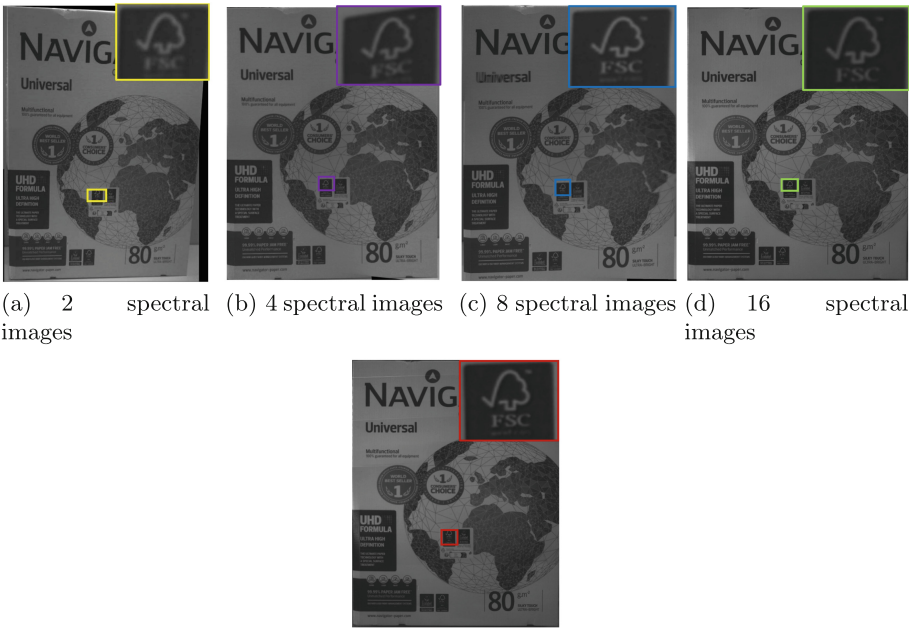
(d) 16 spectral images (e) 32 spectral images

Fig. 5. Spectral Stitched Images of the Same Object, a Calendar: Each stitching composed of various images from different distances—(a) from a distance of 130 cm, (b) from 93 cm, (c) from 76 cm, (d) from 58 cm, (e) from 35 cm.

adjacent images while preserving the spectral information captured by each image.

We assessed the visual appearance of the stitched images across different numbers of images to investigate their impact on image quality. Our analysis aimed to discern whether increasing the number of images enhances the quality of the stitched output or amplifies the probability of distortion. By comparing results obtained from varying numbers of images, we sought to identify trends regarding the influence of image count on image quality. Additionally, we examined the effects on image quality metrics based on natural image statistics such as BRISQUE and NIQUE to determine whether increasing the number of images correlates with improved objective measures of image quality. This investigation aimed to provide insights into the optimization of spectral imaging techniques, particularly in understanding the trade-offs between the number of images used and the resulting image quality.

Initially, it might seem intuitive that moving closer to the object and capturing more images would inherently lead to superior stitching results. However, our experiments challenge this assumption. We observed that as the distance between the camera and the object decreased and the number of images increased, the likelihood of encountering distortion and high-frequency blur in the stitched images also increased (see Fig. 7). This phenomenon suggests that proximity and image abundance do not guarantee improved stitching outcomes; instead, they may exacerbate issues related to image alignment, blending, and, notably, focus.



(a) 2 spectral images (b) 4 spectral images (c) 8 spectral images (d) 16 spectral images

(e) 32 spectral images

Fig. 6. Spectral Stitched Images of the Same Object, a Calendar: Each stitching composed of various images from different distances: (a) from a distance of 130 cm, (b) from 93 cm, (c) from 76 cm, (d) from 58 cm, (e) from 35 cm.

Indeed, one significant factor contributing to reduced stitching quality at closer distances is the increased risk of the camera losing focus. As the camera moves closer to the object, achieving and maintaining optimal focus becomes increasingly challenging, particularly when capturing multiple images to be stitched together. This potential loss of focus further compounds the challenges associated with proximity and image abundance, contributing to the degradation of stitching quality.

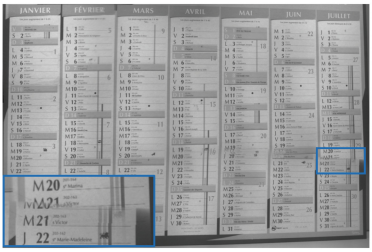
Furthermore, our analysis uncovered a interesting trend regarding the relationship between the number of images used for stitching and the resulting performance metrics and visual quality. While increasing the number of images and decreasing the distance to the object initially led to notable enhancements in stitching quality, we identified a critical threshold beyond which the marginal gains began to diminish. Beyond this threshold, the improvements in performance metrics and visual fidelity stagnated as depicted in the Table 1.

These findings underscore the importance of carefully balancing the trade-offs between proximity, image quantity, and stitching quality in spectral image stitching applications. While closer distances and more images offer the potential for enhanced detail and fidelity, they also introduce challenges related to

Table 1. Average Performance based metrics for various objects.

Nb of Images	BRISQUE ↑	NIQUE ↓
2	32.17	6.07
4	41.81	5.04
8	52.04	3.72
16	62.08	2.49
32	62.84	2.30

focus maintenance that can compromise the final stitching results. Understanding the nuanced dynamics at play in the stitching process is crucial for optimizing performance and achieving the desired balance between image quality and computational efficiency.



(a) 64 spectral images

Fig. 7. Unwanted artifacts appearing. Displayed spectral image at 420 nm.

5 Conclusion

In this paper, we proposed a sequential demosaicking and stitching algorithm for spectral imaging using a hand-held SFA camera from SILIOS Technologies [1]. Our approach involves capturing multiple raw images of a single object and stitching them together to obtain a higher spatial resolution and more detailed images. Through a series of experimentation and analysis, we have demonstrated the efficacy of our algorithm in enhancing the quality of spectral image stitching.

Moreover, Our study shows that capturing spectral images at a close distance to an object results in more images to stitch together, leading to highly detailed panoramas. However, this close proximity also increases the likelihood of distortion problems. Furthermore, The potential applications of this research include fields such as remote sensing, agriculture, and cultural heritage, where having both high spatial and spectral resolution in images is crucial. By leveraging these insights, future advancements in spectral image stitching can be driven towards more robust and reliable solutions that meet the diverse needs of the imaging community.

Acknowledgements. We gratefully acknowledge SILIOS Technologies [1] for generously lending us the CMS-C camera for use in our experiment. The high-quality equipment provided invaluable data and contributed significantly to the success of our research. We are thankful for their support and collaboration.

References

1. SILIOS multispectral cameras CMS series. https://www.silios.com/cms-series. Accessed 01 May 2024
2. Arad, B., Timofte, R., Ben-Shahar, O., Lin, Y.T., Finlayson, G.D.: NTIRE 2020 challenge on spectral reconstruction from an RGB image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 446–447 (2020)
3. Bayer, B.: Color imaging array. United States Patent, no. 3971065 (1976)
4. Bolles, R.C., Fischler, M.A.: A RANSAC-based approach to model fitting and its application to finding cylinders in range data. In: IJCAI, vol. 1981, pp. 637–643 (1981)
5. Brauers, J., Aach, T.: A color filter array based multispectral camera. In: 12. Workshop Farbbildverarbeitung, Ilmenau, pp. 5–6 (2006)
6. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. Int. J. Comput. Vis. **74**, 59–73 (2007)
7. Chini, M., Chiancone, A., Stramondo, S.: Scale object selection through a hierarchical segmentation by a multi-spectral per-pixel classification. Pattern Recogn. Lett. **49**, 214–223 (2014)
8. David, L.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2004)
9. Derpanis, K.G.: The Harris corner detector. York University **2**, 1–2 (2004)
10. Feng, K., Zhao, Y., Chan, J.C.W., Kong, S.G., Zhang, X., Wang, B.: Mosaic convolution-attention network for demosaicing multispectral filter array images. IEEE Trans. Comput. Imaging **7**, 864–878 (2021)
11. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2003)
12. Junior, J.D.D., Backes, A.R., Escarpinati, M.C.: Detection of control points for UAV-multispectral sensed data registration through the combining of feature descriptors. In: VISIGRAPP (4: VISAPP), pp. 444–451 (2019)
13. Lapray, P.J., Wang, X., Thomas, J.B., Gouton, P.: Multispectral filter arrays: recent advances and practical implementation. Sensors **14**(11), 21626–21659 (2014)
14. Lin, C.C., Pankanti, S.U., Natesan Ramamurthy, K., Aravkin, A.Y.: Adaptive as-natural-as-possible image stitching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1155–1163 (2015)
15. Liu, S., Zhang, Y., Chen, J., Lim, K.P., Rahardja, S.: A deep joint network for multispectral demosaicking based on pseudo-panchromatic images. IEEE J. Sel. Top. Sig. Process. **16**(4), 622–635 (2022)
16. Mangai, U.G., Samanta, S., Das, S., Chowdhury, P.R., Varghese, K., Kalra, M.: A hierarchical multi-classifier framework for landform segmentation using multispectral satellite images-a case study over the Indian subcontinent. In: 2010 Fourth Pacific-Rim Symposium on Image and Video Technology, pp. 306–313. IEEE (2010)
17. Mihoubi, S., Losson, O., Mathon, B., Macaire, L.: Multispectral demosaicing using intensity-based spectral correlation. In: 2015 International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 461–466. IEEE (2015)

18. Mihoubi, S., Losson, O., Mathon, B., Macaire, L.: Multispectral demosaicing using pseudo-panchromatic image. IEEE Trans. Comput. Imaging **3**(4), 982–995 (2017)
19. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)
20. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a "completely blind" image quality analyzer. IEEE Sig. Process. Lett. **20**(3), 209–212 (2012)
21. Monno, Y., Kikuchi, S., Tanaka, M., Okutomi, M.: A practical one-shot multispectral imaging system using a single image sensor. IEEE Trans. Image Process. **24**(10), 3048–3059 (2015)
22. Nie, L., Lin, C., Liao, K., Liu, S., Zhao, Y.: Unsupervised deep image stitching: reconstructing stitched features to images. IEEE Trans. Image Process. **30**, 6184–6197 (2021)
23. Pan, Z., Li, B., Bao, Y., Cheng, H.: Deep panchromatic image guided residual interpolation for multispectral image demosaicking. In: 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1–5. IEEE (2019)
24. Pozo, S.D., et al.: Multispectral imaging in cultural heritage conservation. ISPRS - Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **42**, 155–162 (2017)
25. Rathi, V., Goyal, P.: Generic multispectral demosaicking using spectral correlation between spectral bands and pseudo-panchromatic image. Sig. Process. Image Commun. **110**, 116893 (2023)
26. Su, W.H., Sun, D.W.: Multispectral imaging for plant food quality analysis and visualization. Compr. Rev. Food Sci. Food Saf. **17**(1), 220–239 (2018)
27. Wang, X., Thomas, J.B., Hardeberg, J.Y., Gouton, P.: Discrete wavelet transform based multispectral filter array demosaicking. In: 2013 Colour and Visual Computing Symposium (CVCS), pp. 1–6. IEEE (2013)
28. Zaragoza, J., Chin, T.J., Brown, M.S., Suter, D.: As-projective-as-possible image stitching with moving DLT. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2339–2346 (2013)
29. Zhang, Y., Mei, X., Ma, Y., Jiang, X., Peng, Z., Huang, J.: Hyperspectral panoramic image stitching using robust matching and adaptive bundle adjustment. Remote Sens. **14**(16), 4038 (2022)
30. Zhao, B., et al.: PPI edge infused spatial-spectral adaptive residual network for multispectral filter array image demosaicing. IEEE Trans. Geosci. Remote Sens. (2023)
31. Zheng, J., Wang, Y., Wang, H., Li, B., Hu, H.M.: A novel projective-consistent plane based image stitching method. IEEE Trans. Multimedia **21**(10), 2561–2575 (2019)