IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 July 2023, accepted 4 September 2023, date of publication 15 September 2023, date of current version 27 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3315769

APPLIED RESEARCH

Low Pixel Resolution Hyperspectral Image Mosaics Generation Using Learning-Based Feature Matching

CHAMIKA JANITH PERERA[1], (Member, IEEE), CHINTHAKA PREMACHANDRA[2], (Senior Member, IEEE), AND HIROHARU KAWANAKA[1], (Member, IEEE)

[1]Graduate School of Engineering, Mie University, Tsu, Mie 514-8507, Japan
[2]Department of Electronics Engineering, Shibaura Institute of Technology, Tokyo 135-8548, Japan

Corresponding author: Chinthaka Premachandra (chintaka@shibaura-it.ac.jp)

ABSTRACT Remote sensing has become a key component of precision agriculture in the recent decade. Hyperspectral imaging is one key technology that is predicted to be one of the primary decision-making tools in remote sensing-based precision agriculture. Unmanned Aerial Vehicle-based hyperspectral image acquisition is becoming viable due to the reduced cost and the form factor of recently introduced hyperspectral cameras. However, these advantages come at the cost of pixel resolution. Due to factors such as uniform textured surfaces in farmlands, low-pixel resolution, and repeated patterns, traditional stitching methods are unsuccessful at identifying matched features that are needed in generating mosaics from these images. Generating mosaics is a key step in the decision-making process since it opens the ability to interpret information field-wide instead of per-image information interpretations. This paper proposes an image mosaic generation pipeline based on LoFTR - a local feature matching method using transformers as a feature matcher for the low-pixel resolution hyperspectral images. Furthermore, the GPS point-based optimization method is also presented in order to minimize the computational cost and allow the multicore processing capability. The proposed method was evaluated using several field datasets obtained using a low-resolution hyperspectral camera and an unmanned aerial platform. Results present successfully stitched image cubes that could be used in future analysis tasks in agriculture-related decision-making processes.

INDEX TERMS Feature detection, hyperspectral image stitching, mosaic generation, feature matching.

I. INTRODUCTION

The importance of precision agriculture is becoming more and more evident due to various challenges faced by today's farmers. Climate change repercussions have impacted the agriculture industry heavily and losses are ever increasing [1]. Apart from climate change impacts, pest-related losses, soil degradation, and decreasing resource availability such as water also impact towards the losses. Precision agriculture provides the tools that can be used in the decision-making process to mitigate these losses. Spectral sensing is one such tool used in precision agriculture in the decision-making process.

The associate editor coordinating the review of this manuscript and approving it for publication was Tallha Akram.

Out of different spectral imaging methods, multispectral imaging technology can be considered the most widely used in precision agriculture-related remote sensing tasks. Hyperspectral imaging was limited to the research applications due to its high entry cost and the large form factor. However, with the advancement of spectral imaging technologies, both size and cost have come down to a stage where now it's being widely used in UAV-based image acquisition systems [2], [3], [4], [5], [6]. Compared to traditional imaging platforms such as satellite-based acquisition and manned aircraft-based acquisition, UAV-based imaging platforms provide the advantages of high spatial and temporal resolution. From the literature on UAV-based hyperspectral imaging technology, two main image-capturing methods could be identified [2]. PushBroom (Line scanning) sensors [7], [8] where the pixel

line of the image is recorded at one time and the image is generated by combining line-by-line acquisitions. Snapshot sensors [9], [10] where the entire image is captured at once. Snapshot sensors are recently becoming popular due to their advantage of fast acquisition time in UAV-based imaging tasks.

Mosaicing is the process of stitching multiple images into one image [11]. In the remote sensing domain, a mosaic of the captured images is a single image obtained by stitching the images captured using an aerial platform. Interpreting the data and decisions relative to a stitched image is more convenient and user-friendly than interpreting relative to each captured image separately. When the images are interpreted relative to the image-by-image basis, it is difficult to understand the spatial relationship between the image and the actual captured image area. Further, this is especially problematic with low-altitude flying instances with the cameras having narrower optics. Images produced with this combination, such as hyperspectral aerial images with low pixel resolution, gain a great deal from being composed into a mosaic from the captured images.

The typical work process of image mosaic generation includes identifying the features of each image, matching features, calculating the transformation needed to align every image to one coordinate system, and finally blending the images into one by deciding on a strategy to blend overlapping areas without visible seams [12]. However, this process fails at the feature detection and matching stage for hyperspectral images taken from recently introduced low-cost snapshot hyperspectral cameras. This is due to their low pixel resolution and narrower optics not being able to produce enough features needed in transformation calculation accurately. For example, the recently introduced Cubert Ultris 5 camera's resolution is limited to 290px $\times$ 270px. [13] and Field of View is limited to 15°. One method to mitigate this problem is by flying at relatively high altitudes so that the camera captures more features from the scene that could be identified by traditional feature detection methods. However, doing so will reduce the spatial resolution of the image since this will increase the area captured by one pixel. This in turn would result in the loss of the advantage in UAV-based hyperspectral imaging, where leaf-level spectral information can be easily captured by flying at low altitudes.

In this research paper, we investigate the issue of traditional feature-matching methods not being able to identify enough features from low-resolution hyperspectral images and we propose the use of a learning-based feature detection approach to overcome the issue. Based on this feature-matching method, the stitching process is proposed to stitch the low-resolution hyperspectral images to obtain the stitched mosaic. Preliminary results indicate the possibility of using the proposed approach to create hyperspectral mosaics. The main contributions of this paper can be listed as follows.

- Global Positioning System (GPS) information-based image segmenting method: Images are divided into subgroups based on their GPS location information to reduce the computational cost when stitching. This is achieved by reducing the number of matches that need to be considered compared to traditional approaches. Further, it allows the possibility of parallel processing image groups separately.
- Stitching pipeline based on LoFTR (LOcal Feature matching with TRansformers) feature matching: The use of detector-free local feature matching with transformers will be discussed in the context of Low-resolution hyperspectral images. Then the overall methodology followed to obtain the hyperspectral mosaic is presented.

The stitched mosaic using the proposed method was geometrically consistent with the captured target area and spectrally consistent within different targets.

II. RELATED WORK

Image registration is the process of transforming two or more images into a single plane using the common features of the images and it could be identified as a key step in the mosaic generation process. Feature detection-related research work can be divided into two main subsections focusing on image registration of general image stitching research and aerial image-specific stitching research publications. In general image stitching, image registration can be subdivided into two main areas. Intensity-based methods and Feature-based methods [14]. Under intensity-based methods, algorithms such as frequency domain methods [15], spatial domain methods [16], and optical flow methods [17] can be identified. Under feature-based image registration algorithms, geometrical feature identification methods and learning-based feature detection methods can be identified [14]. However, out of these different methods, most image-stitching applications focus on using area-based registration and feature-based registration [18]. This is true for aerial image-specific stitching research as well. 2022 review of existing UAV image mosaicing research [19] indicates that the most widely used methods used in the registration process are Scale-Invariant Feature Transform (SIFT) [16], Speeded-Up Robust Features (SURF) [20], Harries corner detection [21], Features from Accelerated Segment Test (FAST) [22], Binary Robust Invariant Scalable Keypoints (BRISK) [23] and Oriented FAST and Rotated BRIEF (ORB) [24]. All these methods depend on feature-based descriptors that define the point it describes. However, these feature-based detectors perform poorly when there are factors such as poor texture, viewpoint change, illumination variation, repetitive patterns, and motion blur [25]. Poor textures and repetitive patterns are common occurrences in low-resolution hyperspectral images taken using drone platforms. Hence the need arises to use alternative feature-detecting methods to overcome the aforementioned problem.

With the recent advancements in deep learning-based technologies and SLAM (Simultaneous Localization And Mapping) technologies, many learning-based feature-matching methods are being introduced. These learning-based methods can be divided into two subgroups, Detector-based

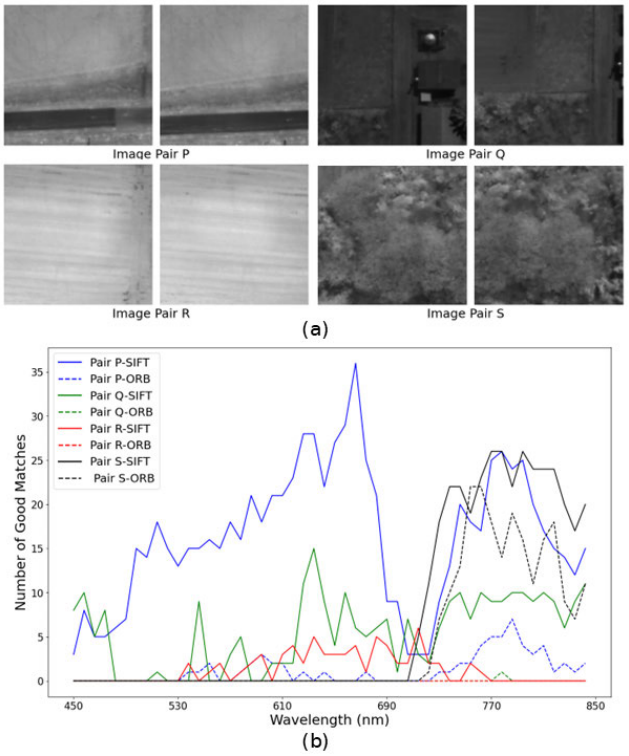

FIGURE 1. Detected feature comparison between SIFT for image pairs P, Q, R, and S.

feature-matching methods, and detector-free feature-matching methods. Several detector-based and learning-based methods, such as Learned Invariant Feature Transform (LIFT) [26], MagicPoint [27], SuperPoint [28], and SuperGlue [29], have demonstrated superior performance in scenarios with changing illumination compared to traditional feature matching algorithms. However, these learning-based methods still face the limitation of feature detectors not being able to identify features that can be easily matched across different images. Especially in the areas of indistinctive regions such as uniform textured areas or low contrast areas.

In order to overcome these issues recent research is focusing on detector-free learning-based feature matching methods. NcNet [30], sparse NcNet [31], and DualRC-Net [32] can be identified as some of the models proposed in this area with promising results. Recently the use of Transformers [33] in vision-related tasks has attracted a lot of focus due to the computational efficiency they promise. LoFTR [25], the feature-matching model incorporated in this proposed stitching method is based on Transformers and the paper indicates that it outperformed all of the existing methods even in problematic illumination changing situations, and uniform textured images. However, LoFTR was neither trained nor tested on low-resolution image data sets, and performance in low-resolution datasets was evaluated during this research.

Hyperspectral image-specific stitching methods are discussed by several authors in the literature. Table 1 summarizes the available literature that we could locate specifically

TABLE 1. Hyperspectral image stitching literature overview.

Paper	Feature Detection Method	Input Image Resolution	Image Acqusition Altitude	Camera Specs
J. Fang et al	Spline Sparse Bundle Adjustment (SSBA)	640px x Length*	N/A	Push-broom
Z. Peng et al	SIFT and mTopKRP	960 x 1057 pixel	300m	Push-broom
Y. Mo et al	SuperPoint & SuperGlue	Several datasets Min. size 640x752 pixels	N/A	Push-Broom
Y. Zhang et al	SuperPoint	960 x 1057 pixels	300m	Push-Broom
L .Yi et al	SIFT	480 x Length*	200m, 100m, 90m	Push-Broom

*Specific mention to the length on the length of the image is not stated in the paper. However, images show more than 3 times the width for the images used.

for hyperspectral image stitching. Fang et al. [34] discusses the usage of the spline Sparse Bundle Adjustment (SSBA) method for image registration. Mo et al. [35] and Zang et al. [36] discuss using SuperPoint [28] and SuperGlue [29] for the feature detection and matching step. Peng et al. [37] and Yi et al. [38] discuss using SIFT as the feature extractor and in the paper [37] authors have used multiscale top K rank preservation (mTopKRP) [39] to obtain a robust set of matches to calculate homography. Hyperspectral images used in all of the papers were above 480 pixels in width. Even though the two papers do not mention the specific length of the image, their figures indicate it is at least more than 3 times the width of the images. All of the papers use images acquired with a pushbroom sensor. Most of the other stitching-related papers such as [40] dive into the specific stitching problems associated with pushbroom sensors which would not be applicable to the images captured using snapshot imaging systems. However, we could not find any existing research articles that focused on low-resolution snapshot hyperspectral image matching and stitching further highlighting the need for identifying suitable methods.

III. PROBLEM DEFINITION

This section further investigates the aforementioned problem in detail. Figure 1 (a) indicates a sample image set of several image acquisition sessions. Each pair represents two adjacent images and images were selected to highlight the feature detection challenges with existing methods. Figure 1 (b) plots the features obtained for each band of the hyperspectral image using SIFT and ORB algorithms. These two were selected because the literature suggests these two are the most used feature detection and matching algorithms in aerial imaging research [19]. Feature matches identified with SIFT are indicated with a solid line and ORB matches are indicated with dotted lines. Image pair P performed well compared to other images with SIFT detection, where most of the

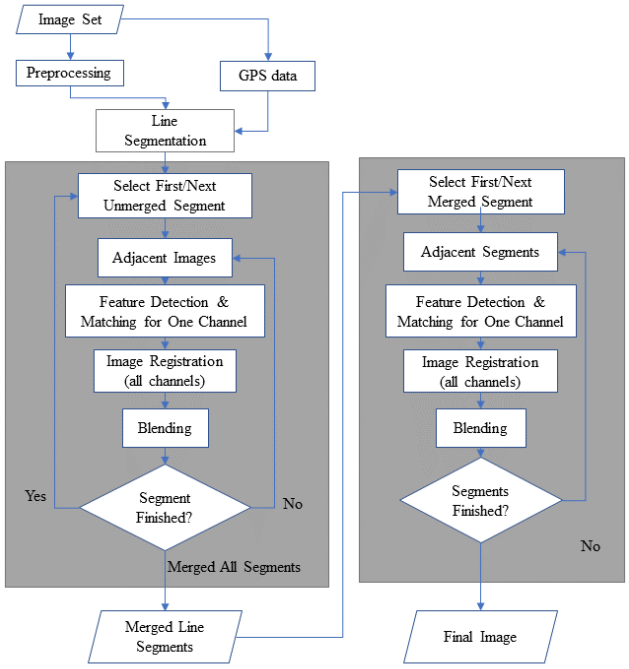

FIGURE 2. Overall methodology followed.

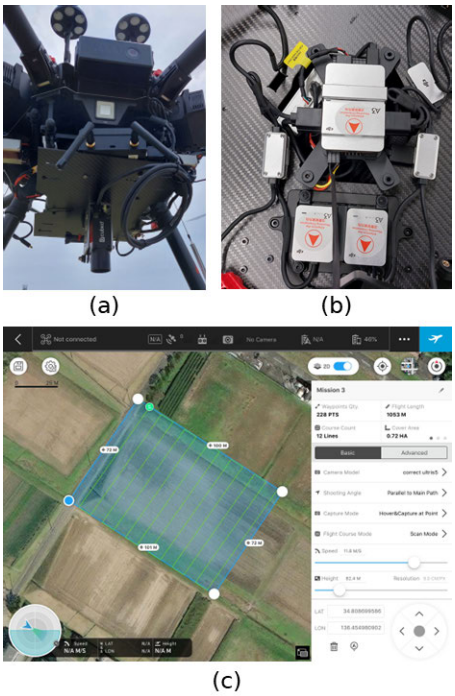

FIGURE 3. (a) Hyperspectral camera mounted on the drone. (b) A3 flight controller. (c) DJI ground station software interface.

bands were capable of producing feature matches of more than 15. However, the other 3 pairs did not perform well with either of the feature detection algorithms. Image pair S did perform well after the 700nm range with both SIFT and ORB detectors. Instead of using single-band feature matching, one other option would be to fuse different bands and obtain features from the fused images such as RGB fused images. However, an evaluation done using all possible 3 band combinations revealed that there were no considerable improvements from that approach either. When the two most poorly performed image pairs were evaluated, the maximum number of matches obtained using pair Q was 21, and pair R produced 36 maximum matches. This is a slight increase from the single-band evaluation. However, the fused bands were completely different from one another, and searching for an ideal band selection for each image by calculating all the combinations is not a viable solution either. Hence a need for a robust feature detection method is highlighted.

IV. METHODOLOGY

The overall methodology followed is depicted in Figure 2. Figure 3 (a) presents the camera mounted on the drone platform. After the image acquisition is carried out each hyperspectral image is pre-processed to convert the recorded digital values to reflectance values. Then the recorded GPS data is used to group the GPS points into lines according to the flight path. Each line segment's images corresponding with each GPS point are loaded into an array sequentially. Adjacent images are used to find the features, calculate the transformation, and then stitch. Then the array is replaced with the stitched images and the process is repeated until there is only one single image in the sequence. The same process is repeated for each image group. Finally, stitched line groups are then processed in the same adjacent feature detection methods to obtain the final image. The following sections will discuss the steps in detail.

A. IMAGE ACQUISITION AND INITIAL PROCESSING

As mentioned previously, the hyperspectral camera used in this research is Cubert Ultris 5 [13]. Table 2 presents the main specifications of the camera. The camera was mounted on a DJI M600 pro [41] drone platform. Mission planning for the data acquisition session was carried out using the DJI Ground Station software [42] presented in Figure 3 (c). Where a shutter signal was generated in the DJI A3 flight controller (Figure 3 (b)) in order to capture an image. The camera was triggered using a PWM relay which was triggered by the aforementioned shutter signal. Captured hyperspectral images are stored in the Pokini F2 Single Board Computer(SBC). The camera software on the SBC records the GPS information along with the image data.

Before every image acquisition session, a calibration session was carried out in order to convert the radiance captured by the camera sensor to reflectance information. First, the camera was turned on till it reached thermal stability. Then a dark image was captured by covering the lens using a lens cap. Later, a white image was captured using the provided calibration card. The black reference image and the white reference image were used to calculate the reflectance. Detailed steps followed including the equipment setup and the calibration steps were discussed in this paper [43]. This whole process of preprocessing the captured images

TABLE 2. Specifications of the camera.

Wavelength Range	450 - 850 nm
Spatial Resolution	290 x 275 pixel
Spectral Bands	51
Spectral Sampling	8 nm
FOV (Field of View)	15°
Weight	126 g
Dimensions	29 x 29 x 65 mm
Bit depth	12 - bit

was carried out using the manufacturer's provided Python Software Development Kit (SDK), which further includes the initial camera-specific calibrations such as geometrical correction and sensor-specific corrections. Processed 16-bit reflectance data were then saved as numpy objects for future processing. The original image captured by the camera is in 12-bit format and the converted 16-bit image is truncated by the SDK to avoid resolution gain. At the same time, GPS information from each file was extracted and saved separately.

All image acquisition sessions were carried out on the Mie University farm located in the Mie prefecture of Japan. Further, all acquisition sessions were carried out between 10.00 AM and 11.30 AM on cloudless clear-skied days. DJI Ground Station [42] was used to plan each mission which calculated the number of images needed to cover the planned mission. Each mission was planned with at least 80% front and side overlaps. Further, the Hover and Capture mode was used in order to reduce the motion blur and to keep the camera axis parallel with the nadir. Mission-specific settings for each dataset such as the altitude are discussed in the results section.

Algorithm 1 GPS Point-Based Image Grouping Algorithm

```
Initialize first segment from GPS_pnts[0, 1]
for i + 2 ← len(GPS_pnts) do
    Calculate θ, α
    Calculate d_current, d_last
    d_diff = d_current − d_last
    θ_diff = θ − α
    if d_diff <= d_tresh or θ_diff <= θ_tresh then
        CurrentSegment ← GPS_pnts[i]
    else
        Start a new segment
        Initialize new Segment with GPS_pnts[i, i + 1]
    end if
end for
```

B. OPTIMIZED STITCHING USING GPS INFORMATION

The traditional method of image stitching to create a mosaic is to first identify features and descriptors from each image and identify matches in between images. Then depending on the matches, the stitching order is determined by grouping the matched images into subsets. A confidence parameter is calculated to check which image is most likely to be adjacent to the next image [44]. However, this matching of each image with other images is a time and resource-consuming process [45]. In order to overcome this issue different GPS-based localization methods are being proposed to optimize the stitching process [18], [44]. In this section, the GPS location-based optimization method is proposed to minimize the processing time of stitching by grouping the images into subgroups where each subgroup can be processed separately to increase computational efficiency.

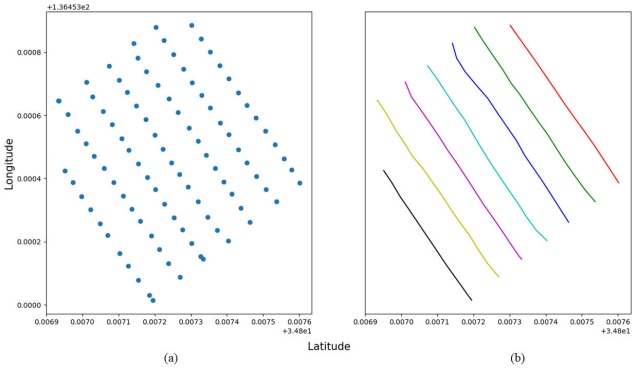

FIGURE 4. (a) GPS points of the recorded images (b) Segmented images subsets.

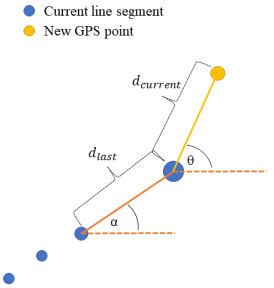

FIGURE 5. Parameters used in identifying line groups.

Algorithm 1 presents the overall procedure followed to group the images into subgroups. This grouping is achieved by identifying lines and the turns that the aircraft moved during the flight and separating each line after each new turn. Images that represent the points of each line are grouped and matched. This algorithm assumes that the flight planning software created the flight path in straight lines and that images are captured in a sequence. First, two images are used to initialize the first segment. From the next GPS point onwards θ: tangent angle between the current point and the last point of the segment and α: tangent angle between the last two points of the current segment is calculated. Further, $d_{current}$: distance between the current point and the last point and d_{last}: distance between the last two points of the current segment is calculated. The aforementioned angles are depicted in figure 5. Angle difference $\theta - \alpha$ and the distance difference $d_{current} - d_{last}$ were calculated next. If the angle difference or the distance difference is below the threshold values set current point is added to the current segment. If not, the current segment will be saved and the new segment will be initialized using the current GPS point and the next one. Figure 4(b) presents the grouped image sequences from a set of GPS points shown in figure 4 (a).

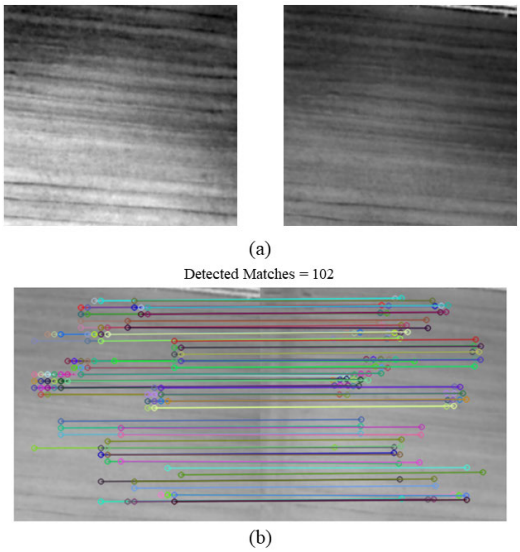

FIGURE 6. (a) Images compared (b) Matches obtained using LoFTR.

C. FEATURE MATCHING, TRANSFORMATION CALCULATION, AND STITCHING

Algorithm 2 Stitching Two Adjacent Images

```
Image1[:,:,Selected_band], Image2[:,:,Selected_band]
Convert to 8-bit images
Find matched Keypoints
if Keypoint_confidence > Threshold then
    Keep Keypoints
end if
Calculate Affine Transformation
if Affine Transformation is None then
    Calculate Homography Transformation
end if
Wrap each band using the Transformation matrix
Blend Each Band
```

Feature detection and matching of the proposed method is carried out by the LoFTR algorithm [25]. LoFTR is a detector-free method that performs local feature matching in image pairs. LoFTR incorporates a detector-free design to extract position and context-dependent local features, which are transformed into feature representations in order to match between images. It uses a Convolutional Neural Network (CNN) to extract multi-level features from the images being compared. A Local Feature Transformer module is then used to extract the position and context-dependent local features. Using linear transformers in the algorithm reduces the computation cost making it suitable for SLAM applications. The use of self-attention and cross-attention layers improve the accuracy of the predictions by allowing the model to attend to important feature and patterns within the input data. Furthermore, the model can handle different types of input data and match them more accurately by using the model's ability to use an optimal transport layer or a dual-softmax operator in establishing coarse-level matches. Figure 6 (a) shows an image pair that when processed with SIFT, SURF, ORB, and BRISK, could not identify any feature matches between images. Figure 6 (b) shows the same two images processed with LoFTR managed to identify a total of 681 feature matches and there were 102 feature matches with 0.9 or higher confidence levels.

Overview of stitching two adjacent images shown in the algorithm 2. Implementation of the LoFTR for the hyperspectral images was carried out using the kornia computer vision library [46]. Since the LoFTR function expects 8-bit images, the selected channel image was converted into 8-bit. Even though this step causes some information loss due to the conversion, it won't affect the final stitched image since two converted images are only used to obtain the matched keypoints. Furthermore, selecting the channel is a decision the user should make considering the captured image area. The contrast between pixels will result in slightly higher performance and will lead to high-quality keypoints. In our experiments, we observed that selecting a channel from 600nm to 750nm will give an advantage over the identified number of keypoints in some image pairs.

LoFTR algorithm outputs matched keypoint pairs along with the confidence of each keypoint pair. It was then filtered with a user-defined confidence threshold. Filtered keypoint pairs are then used to calculate the transformation matrix between the two images. Since the hyperspectral camera is mounted perpendicular to the astronomical horizon during flight, and the images are taken with the hover and capture mode, first the algorithm tries to solve for the affine transformation. Where x' and y' are matched coordinates of x and y and equation was solved to obtain the $a_{11}, \ldots a_{12}$ and t_x, t_y using the equation (1).

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

However, if it fails due to perspective distortions happening due to unstable flight conditions such as windy conditions, homography transformation is calculated using the equation (2) where $h_{i,j}$ represents the homography matrix. Theoretically, homography calculation includes the affine transformation and if the matched points are highly accurate, then the terms h_{31}, h_{32}, h_{33} should become close to 0 in case there are no perspective distortions. However, in practice, it was observed that homography will be highly affected by incorrect feature matches that occur seldom with uniform textures and repeated patterns.

$$\begin{bmatrix} x_i' \\ y_i' \\ w_i' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{2}$$

The calculated transformation matrix is then used to calculate the corners of the output image which is then used to calculate the minimum and maximum x and y coordinates of the output image. The translation matrix is then defined using the minimum x, and y values (x_{min}, y_{min}), and each channel

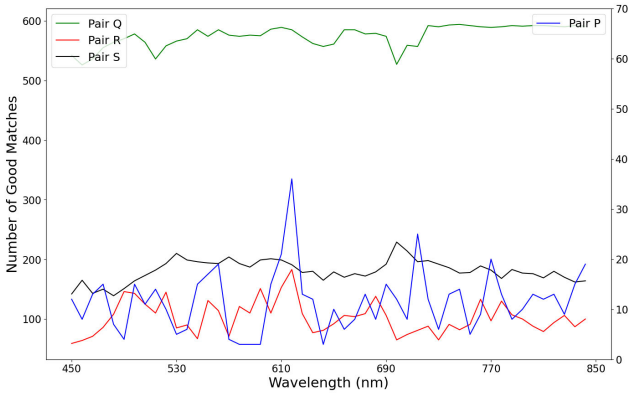

FIGURE 7. Evaluation of the LoFTR feature matching for low-resolution hyperspectral image pairs.

of the second image was warped using this translation matrix presented in equation (3).

$$T = \begin{bmatrix} 1 & 0 & -x_{min} \\ 0 & 1 & -y_{min} \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

If the calculation of the affine transformation was successful each channel of the first image was also warped using the composed matrix of affine and translation matrices. If the affine transformation was unsuccessful, the translation matrix was multiplied by the homography matrix to obtain the full transformation matrix, which was used to warp each channel of image one to the new image plane. Finally, each channel of the warped images was blended separately using a feather blend and stacked to obtain the final stitched image cube with all the channels.

This process was performed on adjacent images until all the images of a line segment were finished. In case there is an odd number of images, The $(n-1)^{th}$ image was repeated in order to make an even number of pairs. Resultant stitched images were again processed using the same strategy and the process was repeated until there was one stitched image of the entire line segment. This strategy also helps in cases where repeating low-resolution features induces false matches between images that are not related. Finally, the same process was repeated for the stitched image groups to obtain the stitched hyperspectral cube.

V. RESULTS AND DISCUSSION

A. EVALUATION OF FEATURE MATCHING METHOD

Figure 7 plots the LoFTR matches obtained for the same images depicted in Figure 1 (a). The number of good matches plotted here is the total sum of matches above a confidence value of 0.9. Pair Q had the highest number of matches with a mean matched number of 574 features. Image pair R produced 102 features as the mean values with the highest value of 183 at band 21 (618nm) and the lowest value of 59 at band 0 (450nm). Image pair S also produced a maximum number of features with 229 matches at band 31 (698nm) and the lowest was 139 at band 0 (450nm). The mean value of the image pair S was 180 matches. Out of all four pairs, Image pair P performed the poorest with a confidence value of 0.9, where only 11 matches were identified as the mean with a maximum of 36 at the 21st band (618nm) and a minimum of 3 matches at band 16 (578). Image pair P's matches are separately plotted on the right axis of Figure 7.

This indicates that there is a possibility for the proposed methods to fail for some images if the user wants to keep a high confidence value of 0.9. Further, it shows that the initial band selection is crucial in a successful stitching operation. As a solution to mitigate the aforementioned issue, an additional modification for the algorithm was implemented to check the number of points with a 0.9 confidence level. If the number of matched points is below a user-defined threshold, it will search for feature matchings in other spectral bands and select the band with the maximum number of features above the threshold. If other spectral bands are unable to produce enough feature matchings, then the algorithm will reduce the confidence level to a user-defined threshold. During the experiments, the second confidence threshold was set to 0.8 and the LoFTR feature detecter always managed to produce enough feature matchings to calculate the transformation parameters.

B. EVALUATION OF STITCHING METHOD

Figure 8 illustrates an experiment carried out using the proposed methodology. This was a small dataset with a total of 54 images captured at an altitude of 50m. At this altitude spatial resolution was 5.5cm/px and the image grouping algorithm divided the captured images into three groups based on their GPS location. Figure 8 (a) presents the intermediate results obtained by the proposed method for each line segment. Figure 8 (b) depicts the final stitched image in both monochrome using band 642nm and the stitched image converted to RGB using the 690nm, 602nm, and 458nm as the red green, and blue bands. the proposed method was successful in producing a final mosaic usable in a decision-making process.

Figure 9 depicts another result of a dataset obtained at a flight height of 120 meters resulting in a 13.1 cm/pixel resolution. A total of 110 images were captured to cover the desired plantation area. Figure 9 (a) presents the RGB converted hyperspectral image and the center plot was sown a few weeks prior for planting for the next season. Hence most of the images taken from the bare soil area contained large amounts of indistinctive regions. However, the proposed method was successful at producing an output geometrically similar to the captured image area. Several sections from this stitched image were then extracted to evaluate the spectral consistency of the stitched image. 20-pixel by 20-pixel regions representing vegetation, soil, asphalt, and region with a mix of weed and soil were selected for evaluation. Selected areas are depicted with bounding boxes in Figure 9 (a). Figure 9 (b) presents the respective spectral plots of the selected regions. It can be observed that the signatures are stable within the bounding box without any abnormal

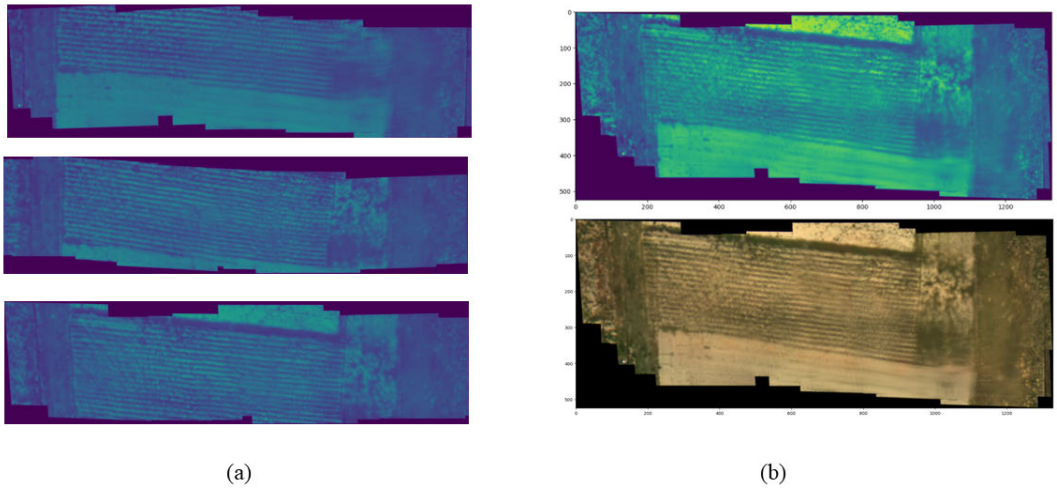

FIGURE 8. (a) Intermediate results of line segment stitchings (b) Final stitched image in monochrome and RGB.

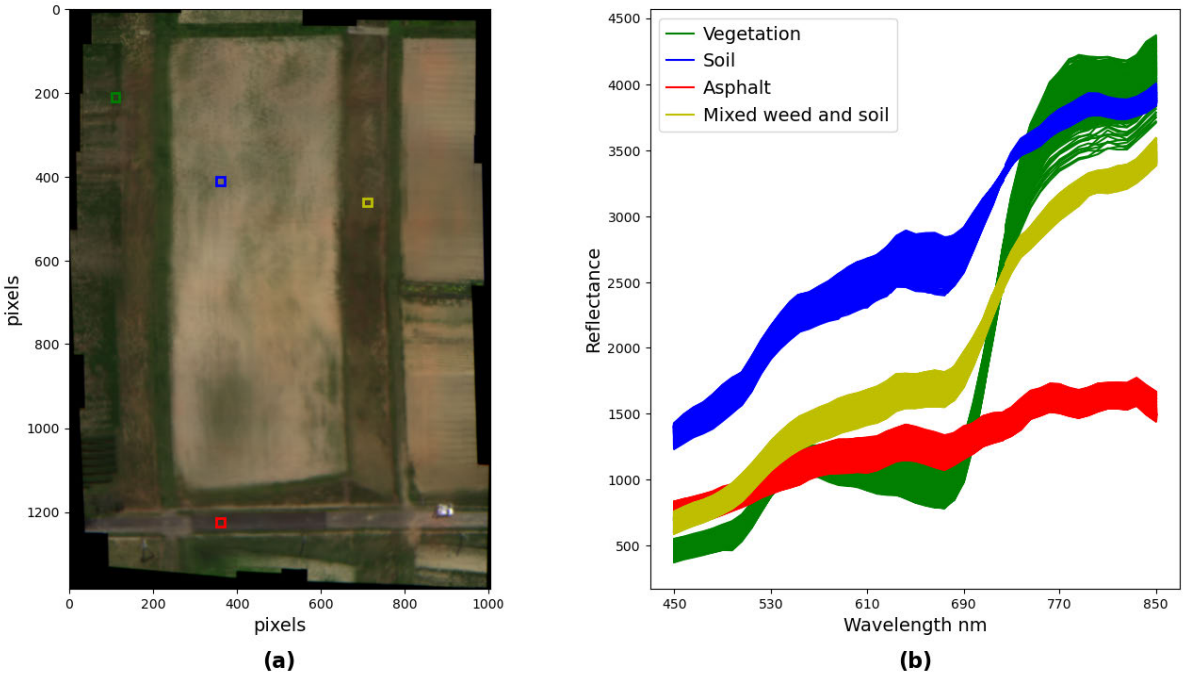

FIGURE 9. (a) Output of the final stitching result in RGB (b) Spectral analysis of different regions of the stitched image.

noise present in each spectral signature. Furthermore, it can be observed that the different targets that were selected, produced unique spectral signatures that could be used to differentiate among different targets. The region selected from the mix of weed and soil indicated signatures in between the soil and the vegetation. Spectral mixing could be identified as the reason for this since, at this spatial resolution leaves and soil will be mixed in the spectral signature of a single pixel.

VI. CONCLUSION AND FUTURE WORKS

In conclusion, this research paper investigates the drawbacks of existing image stitching methods for low-resolution hyperspectral image stitching and proposes a new approach for producing hyperspectral image mosaics that could be used in further decision-making processes. This research paper identified that the existing methods failed at the feature identification and matching steps due to the lack of features in the low-resolution hyperspectral images. A learning-based feature-matching algorithm was incorporated to solve this problem and the initial experiments indicated significant improvements over the traditional feature detection and matching algorithms in use. A stitching pipeline was proposed incorporating the identified feature-matching algorithm and results were presented. Initial results indicate successful mosaic generation from the input images with the limitations of some geometrical distortions and several

ghosting effects around targets such as buildings and structures. However, it's a significant step forward from the traditional methods where it failed to produce any outputs. Furthermore, spectral consistency and differentiability evaluations indicate that the agriculture vegetation areas can be used in further decision-making tasks such as segmentation, classification, activity detection, and generating scene graphs [47], [48].

The incorporated LoFTR algorithm [25] was used with its pre-trained model weights where there were no low-resolution aerial images. Performing a transfer learning step to train with a low-resolution aerial dataset would have the potential to improve the current drawbacks in terms of identifying features. However, since the algorithms used depth maps for training data, the generation of depth maps from hyperspectral images should be taken into consideration when retraining the algorithm. Methods such as zero-shot learning with generative adversarial networks [49] will be investigated in the future to retrain the LoFTR model with hyperspectral depth maps. Furthermore, improving the blending method to get rid of ghosting can be considered as another future work. A model like an optimum seam detection for blending [50] of objects can be a viable solution to improve the stitched images.

REFERENCES

[1] *The Impact of Disasters and Crises on Agriculture and Food Security: 2021*, The Food and Agriculture Organization of the United Nations, Rome, Italy, 2018.
[2] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. Sousa, "Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry," *Remote Sens.*, vol. 9, no. 11, p. 1110, Oct. 2017.
[3] L. W. Kuswidiyanto, H.-H. Noh, and X. Han, "Plant disease diagnosis using deep learning based on aerial hyperspectral images: A review," *Remote Sens.*, vol. 14, no. 23, p. 6031, Nov. 2022.
[4] X. Lyu, X. Li, D. Dang, H. Dou, K. Wang, and A. Lou, "Unmanned aerial vehicle (UAV) remote sensing in grassland ecosystem monitoring: A systematic review," *Remote Sens.*, vol. 14, no. 5, p. 1096, Feb. 2022.
[5] S. Ecke, J. Dempewolf, J. Frey, A. Schwaller, E. Endres, H.-J. Klemmt, D. Tiede, and T. Seifert, "UAV-based forest health monitoring: A systematic review," *Remote Sens.*, vol. 14, no. 13, p. 3205, Jul. 2022.
[6] N. Amarasingam, A. S. A. Salgadoe, K. Powell, L. F. Gonzalez, and S. Natarajan, "A review of UAV platforms, sensors, and applications for monitoring of sugarcane crops," *Remote Sens. Appl., Soc. Environ.*, vol. 26, Apr. 2022, Art. no. 100712.
[7] *VIS NIR Hyperspectral Imagers | BaySpec Bayspec.com*. Accessed: Apr. 3, 2023. [Online]. Available: https://www.bayspec.com/products/vis-nir-hyperspectral-camera/
[8] *VNIR 400 1000nm-Headwallphotonics.Com*. Accessed: Apr. 3, 2023. [Online]. Available: https://www.headwallphotonics.com/products/vnir-400-1000nm
[9] *Hyperspectral Camera for Agriculture and Forestry | HSC-2 | MosaicMill | Rikola Mosaicmill.Com*. Accessed: Apr. 3, 2023. [Online]. Available: https://www.mosaicmill.com/products_other/hs_camera.html
[10] *HSC-2 Hyperspectral Camera 450-800nm Senop Senop.fi*. Accessed: Apr. 3, 2023. [Online]. Available: https://senop.fi/product/hsc-2-hyperspectral-camera-450-800nm/
[11] D. Capel, "Image mosaicing," in *Image Mosaicing Super-Resolution*. Cham, Switzerland: Springer, 2004, pp. 47–79.
[12] Z. Wang and Z. Yang, "Review on image-stitching techniques," *Multimedia Syst.*, vol. 26, no. 4, pp. 413–430, Aug. 2020.
[13] *Gmbh Real-Time Spectral Imaging*. cubert. Accessed: Nov. 15, 2022. [Online]. Available: https://www.cubert-hyperspectral.com/products/ultris-5
[14] R. Feng, H. Shen, J. Bai, and X. Li, "Advances and opportunities in remote sensing image geometric registration: A systematic review of state-of-the-art approaches and future research directions," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 4, pp. 120–142, Dec. 2021.
[15] Y. Li, J. Wang, and K. Yao, "Modified phase correlation algorithm for image registration based on pyramid," *Alexandria Eng. J.*, vol. 61, no. 1, pp. 709–718, Jan. 2022.
[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[17] H. Lee, S. Lee, and O. Choi, "Improved method on image stitching based on optical flow algorithm," *Int. J. Eng. Bus. Manage.*, vol. 12, Jan. 2020, Art. no. 184797902098092.
[18] A. Moussa and N. El-Sheimy, "A fast approach for stitching of aerial images," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 41, pp. 769–774, Jun. 2016.
[19] J. K. Gómez-Reyes, J. P. Benítez-Rangel, L. A. Morales-Hernández, E. Resendiz-Ochoa, and K. A. Camarillo-Gomez, "Image mosaicing applied on UAVs survey," *Appl. Sci.*, vol. 12, no. 5, p. 2729, 2022.
[20] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.
[21] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, 1988, p. 5244.
[22] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Oct. 2005, pp. 1508–1515.
[23] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2548–2555.
[24] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
[25] J. Sun, Z. Shen, Y. Wang, H. Bao, and X. Zhou, "LoFTR: Detector-free local feature matching with transformers," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 8918–8927.
[26] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "LIFT: Learned invariant feature transform," in *Proc. Eur. Conf. Comput. Vis.* Amsterdam, The Netherlands: Springer, Sep. 2016, pp. 467–483.
[27] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Toward geometric deep SLAM," 2017, *arXiv:1707.07410*.
[28] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-supervised interest point detection and description," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 337–33712.
[29] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning feature matching with graph neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 4937–4946.
[30] I. Rocco, M. Cimpoi, R. Arandjelovic, A. Torii, T. Pajdla, and J. Sivic, "NCNet: Neighbourhood consensus networks for estimating image correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 2, pp. 1020–1034, Feb. 2022.
[31] I. Rocco, R. Arandjelović, and J. Sivic, "Efficient neighbourhood consensus networks via submanifold sparse convolutions," in *Proc. IEEE Eur. Conf. Comput. Vis.* Glasgow, U.K.: Springer, Aug. 2020, pp. 605–621.
[32] X. Li, K. Han, S. Li, and V. Prisacariu, "Dual-resolution correspondence networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 17346–17357.
[33] N. Kitaev, Ł. Kaiser, and A. Levskaya, "Reformer: The efficient transformer," 2020, *arXiv:2001.04451*.
[34] J. Fang, X. Wang, T. Zhu, X. Liu, X. Zhang, and D. Zhao, "A novel mosaic method for UAV-based hyperspectral images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 9220–9223.
[35] Y. Mo, X. Kang, P. Duan, and S. Li, "A robust UAV hyperspectral image stitching method based on deep feature matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5517514.
[36] Y. Zhang, X. Mei, Y. Ma, X. Jiang, Z. Peng, and J. Huang, "Hyperspectral panoramic image stitching using robust matching and adaptive bundle adjustment," *Remote Sens.*, vol. 14, no. 16, p. 4038, Aug. 2022.
[37] Z. Peng, Y. Ma, X. Mei, J. Huang, and F. Fan, "Hyperspectral image stitching via optimal seamline detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[38] L. Yi, J. M. Chen, G. Zhang, X. Xu, X. Ming, and W. Guo, "Seamless mosaicking of UAV-based push-broom hyperspectral images for environment monitoring," *Remote Sens.*, vol. 13, no. 22, p. 4720, Nov. 2021.
[39] X. Jiang, J. Jiang, A. Fan, Z. Wang, and J. Ma, "Multiscale locality and rank preservation for robust feature matching of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6462–6472, Sep. 2019.
[40] J. Li, L. Ma, Y. Fan, N. Wang, K. Duan, Q. Han, X. Zhang, G. Su, C. Li, and L. Tang, "An image stitching method for airborne wide-swath HyperSpectral imaging system equipped with multiple imagers," *Remote Sens.*, vol. 13, no. 5, p. 1001, Mar. 2021.
[41] *Matrice 600 Pro Product Information DJI Dji.com*. Accessed: Mar. 17, 2023. [Online]. Available: https://www.dji.com/jp/matrice600-pro/info
[42] *DJI GS Pro DJI Dji.com*. Accessed: Apr. 24, 2023. [Online]. Available: https://www.dji.com/ground-station-pro
[43] C. J. Perera, C. Premachandra, and H. Kawanaka, "Comparison of light weight hyperspectral camera spectral signatures with field spectral signatures for agricultural applications," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2023, pp. 1–3.
[44] L. A. Weber and D. Schenk, "Automatic merging of split construction plans of hydraulic structures," *Engineering*, vol. 99, no. 5, pp. 330–340, 2022.
[45] R. Xiang, M. Sun, C. Jiang, L. Liu, H. Zheng, and X. Li, "A method of fast mosaic for massive UAV images," in *Proc. SPIE*, vol. 9260, pp. 725–733, Nov. 2014.
[46] E. Riba, D. Mishkin, D. Ponsa, E. Rublee, and G. Bradski, "Kornia: An open source differentiable computer vision library for PyTorch," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 3674–3683.
[47] X. Chang, P. Ren, P. Xu, Z. Li, X. Chen, and A. Hauptmann, "A comprehensive survey of scene graphs: Generation and application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 1–26, Jan. 2023.
[48] L. Zhang, X. Chang, J. Liu, M. Luo, Z. Li, L. Yao, and A. Hauptmann, "TN-ZSTAD: Transferable network for zero-shot temporal activity detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 3, pp. 3848–3861, Mar. 2023.
[49] M. Li, P.-Y. Huang, X. Chang, J. Hu, Y. Yang, and A. Hauptmann, "Video pivoting unsupervised multi-modal machine translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 3, pp. 3918–3932, Mar. 2023.
[50] J. Gao, Y. Li, T.-J. Chin, and M. S. Brown, "Seam-driven image stitching," in *Proc. Eurographics (Short Papers)*, 2013, pp. 45–48.

CHAMIKA JANITH PERERA (Member, IEEE) received the B.Sc. and M.Sc. degrees from the University of Moratuwa, Sri Lanka, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree with Mie University, Japan. He was a Research Assistant with the Mechanical Engineering Department, University of Moratuwa, from 2016 to 2018. He was a Research Engineer with the Center for Advanced Robotics, University of Moratuwa, from 2018 to 2019. He finished a short-term research internship with the Shibaura Institute of Technology, Japan, in 2017, and was a Research collaborator for research carried out with The University of New Mexico and Lady Ridgeway Hospital for Children, Sri Lanka, in 2018. Initially, he carried out research in the areas of brain machine interfacing and robotics. However, he shifted his research interests toward UAV-based agriculture sensing for the Ph.D. His current research interests include hyperspectral imaging, machine learning, signal processing, unmanned aerial vehicles, and agriculture robotics. He received the OpenBCI Hackathon First Prize, in 2017, the second place in the Matlab Mini Drone Competition, in 2018, and the Best Paper Award from the MERCon 2017 Conference, in 2017.

CHINTHAKA PREMACHANDRA (Senior Member, IEEE) was born in Sri Lanka. He received the B.Sc. and M.Sc. degrees from Mie University, Tsu, Japan, in 2006 and 2008, respectively, and the Ph.D. degree from Nagoya University, Nagoya, Japan, in 2011.

From 2012 to 2015, he was an Assistant Professor with the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan. From 2016 to 2017, he was an Assistant Professor with the Department of Electronic Engineering, School of Engineering, Shibaura Institute of Technology, Tokyo, where he was an Associate Professor, from 2018 to 2022. In 2022, he was promoted to a Professor with the Department of Electronic Engineering, Graduate School of Engineering, Shibaura Institute of Technology, where he is currently the Manager of the Image Processing and Robotic Laboratory. His research interests include AI, UAV, image processing, audio processing, intelligent transport systems (ITS), and mobile robotics. He is a member of IEICE, Japan; SICE, Japan; RSJ, Japan; and SOFT, Japan. He received the IEEE Sensors Letters Best Paper Award from the IEEE Sensors Council, in 2022; the IEEE Japan Medal from the IEEE Tokyo Section, in 2022; and the FIT Best Paper Award and the FIT Young Researchers Award from IEICE and IPSJ, Japan, in 2009 and 2010, respectively. He served as a steering committee member and an editor for many international conferences and journals. He is the Founding Chair of the International Conference on Image Processing and Robotics (ICIPRoB), which is technically co-sponsored by IEEE. He is serving as an Associate Editor for IEEE Robotics and Automation Letters (R-AL) and *IEICE Transactions on Information and Systems*.

HIROHARU KAWANAKA (Member, IEEE) received the Ph.D. degree in engineering from the Graduate School of Engineering, Mie University, in 2004, and the Ph.D. degree in medical science from the Graduate School of Medicine, Mie University, in 2009. In 2004, he established a company for medical information systems to develop a clinical data warehouse and medical accounting system for national university hospitals. He worked for this company as the Director, from 2004 to 2021. In 2006, he joined Mie University, as an Assistant Professor, where he has been an Associate Professor with the Graduate School of Engineering, since 2017. He was a Visiting Associate Professor with the Suzuka University of Medical Science, from 2017 to 2021. He was also an Assistant to the President, from 2018 to 2023. His current research interests include medical informatics, medical document image analysis, graphics recognition, welfare information systems, ergonomics, evolutionary computations, and their application. He is a member of the SMC Society, the Computer Society, HIMSS, and some Japanese academic societies.