An Efficient Method for Generating UAV-Based Hyperspectral Mosaics Using Push-Broom Sensors

Juan M. Jurado, Luís Pádua, Jonas Hruška, Francisco R. Feito, and Joaquim J. Sousa

***Abstract*—Hyperspectral sensors mounted in unmanned aerial vehicles offer new opportunities to explore high-resolution multitemporal spectral analysis in remote sensing applications. Nevertheless, the use of hyperspectral data still poses challenges mainly in postprocessing to correct from high geometric deformation of images. In general, the acquisition of high-quality hyperspectral imagery is achieved through a time-consuming and complex processing workflow. However, this effort is mandatory when using hyperspectral imagery in a multisensor data fusion perspective, such as with thermal infrared imagery or photogrammetric point clouds. Push-broom hyperspectral sensors provide high spectral resolution data, but its scanning acquisition architecture imposes more challenges to create geometrically accurate mosaics from multiple hyperspectral swaths. In this article, an efficient method is presented to correct geometrical distortions on hyperspectral swaths from push-broom sensors by aligning them with an RGB photogrammetric orthophoto mosaic. The proposed method is based on an iterative approach to align hyperspectral swaths with an RGB photogrammetric orthophoto mosaic. Using as input preprocessed hyperspectral swaths, apart from the need of introducing some control points, the workflow is fully automatic and consists of: adaptive swath subdivision into multiple fragments; detection of significant image features; estimation of valid matches between individual swaths and the RGB orthophoto mosaic; and calculation of the best geometric transformation model to the retrieved matches. As a result, geometrical distortions of hyperspectral swaths are corrected and an orthomosaic is generated. This methodology provides an expedite solution able to produce a hyperspectral mosaic with an accuracy ranging from two to five times the ground sampling distance of the high-resolution RGB orthophoto mosaic, enabling the hyperspectral data integration with data from other sensors for multiple applications.**

***Index Terms*—Hyperspectral imaging, mosaicking, push-broom sensor, unmanned aerial vehicles (UAVs).**

Manuscript received April 9, 2021; revised May 13, 2021 and June 7, 2021; accepted June 9, 2021. Date of publication June 14, 2021; date of current version July 14, 2021. This work was supported in part by the Ministry of Science and Innovation of Spain and the European Union (via ERDF funds) under Grant TIN2017-84968-R in part by the FCT-Portuguese Foundation for Science under Grant UIDB/50014/2020 and in part by the National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020. (*Corresponding author: Juan M. Jurado*.)

Juan M. Jurado and Francisco R. Feito are with the Computer Graphics and Geomatics Group of Jaén, University of Jaén, 23071 Jaén, Spain (e-mail: jjurado@ujaen.es; ffeito@ujaen.es).

Luís Pádua and Joaquim J. Sousa are with the Engineering Department, School of Science and Technology, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal, and also with the Centre for Robotics in Industry and Intelligent Systems, INESC Technology and Science, 4200-465 Porto, Portugal (e-mail: luispadua@utad.pt; jjsousa@utad.pt).

Jonas Hruška is with the Engineering Department, School of Science and Technology,University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal (e-mail: h.jonas@seznam.cz).

Digital Object Identifier 10.1109/JSTARS.2021.3088945

I. Introduction

HYPERSPECTRAL imaging consists of a type of spectral imaging and as such it results from integrating conventional images with spectroscopy methods, enabling gathering both spatial and spectral information from an object [1]. This technique was originally proposed by Goetz *et al.* [1] to be used in earth remote sensing applications. While a common frame camera captures light from the three visible region wavelengths [red, green, and blue (RGB)], a hyperspectral imagery covers a wide variety of bands that go beyond RGB. Hence, each pixel belonging to a hyperspectral image contains information related to hundreds of bands from different regions of the electromagnetic spectrum (generating the so-called hypercube).

Spectral imaging technology was originally used in earth remote sensing applications, mainly in aerial surveillance. It represented a true revolution in satellite-based remote sensing, allowing first the acquisition of multispectral images—a group of few bands belonging to the visible and near-infrared (VNIR) spectral region.

Recent technological developments have enabled the development of a platform capable of overcoming the major issues associated with both manned aircraft and satellites while simultaneously improving spectral and spatial resolutions. Unmanned aircraft vehicles (UAVs) have emerged due to their operational flexibility and accessibility to users [2], allowing remote sensing to gain a new platform capable of better understanding earth's system phenomena [2], [3]. Together with specialized sensors, UAVs are becoming powerful monitoring systems [3] that complement traditional monitoring techniques instead of competing with them [4]. Nowadays, monitoring and data collection systems consist of a combination of different observations/measurements: ground-based measurements, airborne and satellite sensors. Taking advantage of the aforementioned capabilities, UAVs and sensor technologies advances, have stimulated their use in a vast domain of applications [3], [5]–[13]. On the other hand, recent hyperspectral sensors, especially covering wavelengths ranging from 400 to 1000 nm (RGB + VNIR) have evolved regarding a significant reduction of the weight and size, allowing them to be coupled to UAVs [4].

Although the level of maturity reached by UAV-based technology in terms of sensing capabilities [4], the use of hyperspectral imaging systems is still challenging. Nonetheless, the potential inferred by published studies has sparked increasing interest in aerial spectral imaging, leading several companies to invest in the development of UAV-based hyperspectral sensors [4].

However, in addition to the complexity associated with the interpretation of wealth of data, the complexity associated with hyperspectral imagery acquisition and processing remains an important obstacle to its use. In particular, to fully exploit the advantages given by hyperspectral imagery it is crucial to ensure the positional accuracy. Indeed, the great radiometric diversity provided by hyperspectral sensing systems will only be useful if it is geometrically accurate, allowing quantitative analysis to be performed.

Orthophoto maps (ortho mosaics) are one of the basic and most important products generated by UAV imagery, forming also the basis of hyperspectral systems [5]. Readers are referred to [14] to review the most important achievements and challenges on image mosaicking. Different types of hyperspectral sensor configurations are available, including point (whiskbroom), push-broom (line scanning), and snapshot (two-dimensional (2-D) spectral images). The push-broom hyperspectral solutions are very popular due to their high spectral and spatial resolutions, and have been used for a long time by earth observation satellite systems [6]–[10], [15]. This approach can simultaneously acquire a range of spatial information, as well as spectral information corresponding to each spatial point in the scanned range. Images with a spatial dimension (y-line) and a spectral dimension (λ) can be obtained at the same time with a charge-coupled device. Usually, several scan lines are associated in frames (images) that will be stitched together [11], creating a hyperspectral swath. As a result, individual images have a positional accuracy on the order of a few centimeters; however, the resulting orthomosaic will accumulate the positional errors from the stitching process, resulting in a final accuracy in the decimeter/meter range. A scanline sensor is highly sensitive to motion, as such, a stabilized gimbal has to be used to reduce the geometric noise in acquired spectral data, caused by the UAV movements. Every change in pitch, roll, and yaw automatically done to follow the preprogrammed flight plan is transcribed into the acquired data through the use of global navigation satellite systems (GNSS) and inertial measurement unit (IMU) [5]. Thus, and despite the high spectral and spatial resolution offered by push-broom sensors, the spatial accuracy of each scanline is highly dependent on flying conditions, being the final error depending on the GNSS/IMU sensors accuracy [12] and the stability provided by the gimbal.

A review of the published studies based on the use of UAV-based push-broom hyperspectral systems allowed us to conclude that, although widely used, the geometric rectification is still challenging [4]. Zarco-Tejada *et al.* [6] used a Headwall Micro-Hyperspec VNIR push-broom sensor to investigate the early detection of plant diseases. They apply a usual orthorectification procedure based on GNSS/IMU parameters and a digital elevation model (DEM) using PARGE software [16]. Other authors used a dense network of ground control points (GCP) in addition to PARGE software for orthorectification [7], [8]. Globally, a 5 cm root-mean-square error (RMSE) was achieved, using hyperspectral imagery of 2 – 4 cm ground sample distance (GSD).

More recently, and to increase the geometric rectification, light detection and ranging (LiDAR) sensors, known as one of the most accurate ways for geometric data acquisition, were coupled with the hyperspectral sensor. This way, both datasets can be synchronized, mitigating the fact that the UAV platforms are often inaccurate with respect to the sensor resolution, and the platform is also unstable when flying. For example, Sankey *et al.* [9] used a Headwall's SpectralView software (Headwall Photonics, Inc., Bolton, MA, USA) to process the individual hyperspectral frames and to produce the final mosaic, achieving an RMSE of 0.94 m. Despite the overall good results provided by the LiDAR/hyperspectral sensors integration, LiDAR sensors, even those considered low-cost, are still very expensive and require higher payloads. Therefore, this integration setup is not yet accessible to everyone as UAV-based photogrammetric mapping systems. In order to reduce ground sampling efforts, other approaches have been explored. One of the most obvious approaches consists of the use of an RGB sensor to acquire visible and hyperspectral data simultaneously. Thus photogrammetry-based computer vision approaches and structure from motion algorithms are used to create more precise DEM to feed PARGE or SpectralView software. This was done by Ramirez-Paredes *et al.* [17] and Habib *et al.* [18], reaching submetric accuracies.

Finally, popular computer vision algorithms were also tested in coregistration approaches [19]–[22]. Scale-invariant feature transform (SIFT), speeded up robust features (SURF), features from accelerated segment test (FAST), and binary robust independent elementary features (BRIEF) have been used to perform a robust feature detection of key-points between adjacent frames. To explore the most recent developments in different domains, Angel *et al.* [22] proposed a fully automatic workflow to produce highly accurate georectified UAV-based hyperspectral mosaics collected by push-broom sensors, requiring a small number of GCPs. The Headwall Nano-Hyperspec push-broom sensor was used over two experimental crop sites. In both cases, a high-resolution RGB orthophoto mosaic was used (respectively, 0.7 cm GSD and 6 cm GSD). Then, the SURF algorithm was used to extract a set of common features between the RGB reference image and each hyperspectral swath. It was concluded that different illumination conditions and the use of different sensors (RGB/hyperspectral), resulting in distinct geometries, influencing the number of key-points detected and thus the rectification performance. Although the excellent results globally provided by this automatic method, it should be noted that a high computation power was used (200 GB of RAM and CPU with 20 cores) and some hours were needed to conclude the procedure. This is a great advance compared to previous approaches ([5], [6], [11], [16], [18], [23]), despite that significant resources demanded may prevent its broader use. Fang *et al.* [24] used a spline function mosaic method based on bundle adjustment for mosaicking hyperspectral data acquired with a Headwall MicroHyperspect push-broom sensor. An RGB orthomosaic, generated from images obtained simultaneously with hyperspectral data, is used to identify at least ten key-points. The spline function is used to bend the input image so that the image can connect all the key-points at the same time. The method divides the image by the size of the spline function, and the size of the block is modified so that the number of feature points in each block is approximately the same. This process needs to solve the best fitting function iteratively, and the computational

complexity is high. The method produces hyperspectral mosaics with an overall accuracy of about 3 px; however, the authors did not provide the spatial resolution (GSD).

In the scope of the main issues highlighted in the previous paragraph, in this study an automatic workflow to improve the orthorectification of the UAV-based hyperspectral data is presented. A push-broom sensor was used to collect hyperspectral data over a diversified study area to assess the applicability and the effectiveness of the proposed method.

This article is organized as follows. Section II describes the studied area, presents the two aerial platforms used for data acquisition, and describes the experiment design in terms of flight plans and GCP. This section ends with an explanation on the raw hyperspectral data preprocessing; Section III presents, step-by-step, the proposed method and describes the validation setup; Section IV presents the results; Section V discusses the main findings and compares them to the state of the art. Finally, Section VI conlcudes this article.

II. Materials

A. Study Area

This study was designed to evaluate the proposed methodology in most hyperspectral applications. A 4-ha area, in the University of Trás-os-Montes e Alto Douro campus [Fig. 1(a)], Vila Real, Portugal, was selected as a study area [Fig. 1(b)]. This area is representative of the majority of hyperspectral use cases scenarios, as it is formed by a diversity of environments. The southernmost sector of the study area is composed of dense vegetation, in which, due to its homogeneity, it is difficult to identify key-points. The northern sector is mainly occupied by man-made linear features, which facilitate the identification of key-points. In the remaining sectors, there is a mixture of undergrowth and linear entities, such as roads and parking lots. Regarding the topography, the study area is characterized by a medium altitude (500 m in average), with a maximum variation of about 30 m. In this way, hyperspectral swaths with different types of occupation and geometric distortions will be obtained, allowing to evaluate the performance of the developed methodology.

B. Sensors and Platforms

Two different unmanned aerial systems were used for the data collection process. The DJI Matrice 600 Pro (M600) hexacopter was the flight platform used to carry Headwall's Nano-Hyperspec sensor [see Fig. 2(a)]. The sensor is mounted in a Ronin-MX gimbal attached to the M600. This ensures sensor's stability which along with a constant UAV flight speed and trajectory must be assured for an optimal hyperspectral push-broom data acquisition process with no distortions [25]. A 12-mm lens with a horizontal field of view (HFoV) of 21.1° was used. Each line of pixels comprises 640 spatial pixels and 270 spectral bands. The sensor acquires 12-b data, i.e., 4096 brightness levels across the VNIR spectrum, ranging from 400 to 1000 nm, with a sampling interval of 2.2 nm and a full-width at half-maximum of approximately 6 nm [26]. Five global positioning system antennas are mounted on the top of the M600:

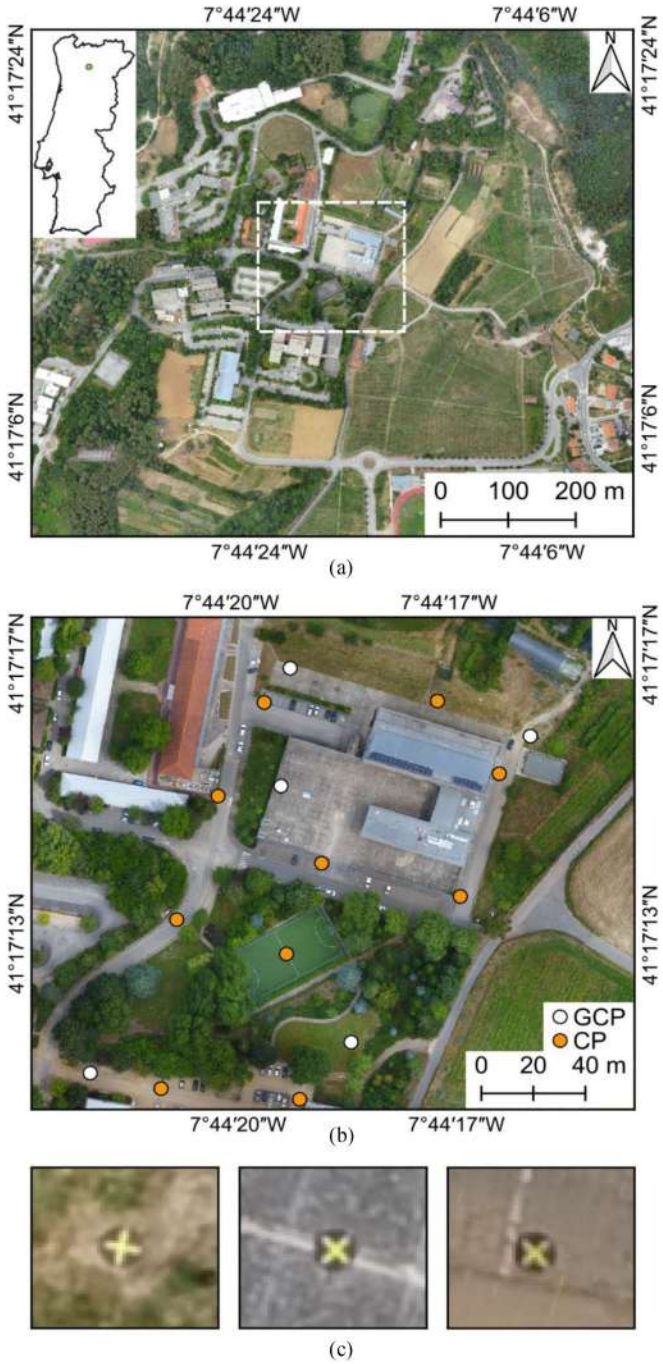
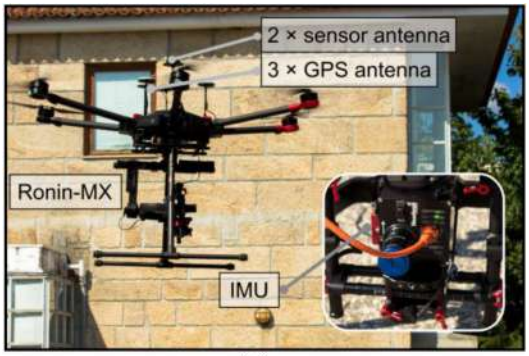


Fig. 1. Study area location. (a) General overview. (b) Detailed visualization of the study area with the locations of the GCP and check points (CP) used for the geocoding and geometric quality assessment of the RGB-orthophoto mosaic. (c) Zoom on the targets used as GCPs.

three—mounted in the upper plat—for the UAV's navigation and attitude determination; and two—mounted in the arms—used for the positioning of the hyperspectral sensor. An IMU is also used to account for the effects of roll, pitch, and yaw in the sensor. With this equipment profile, the M600 has an autonomy up to 25 min, depending on the wind conditions. RGB imagery was acquired using a DJI Phantom 4 quadcopter, which is paired with a three-axis electronic gimbal for camera stabilization. An



(a)

(b)

Fig. 2. Overview of the UAVs used in this study. (a) Matrice 600 Pro and the Headwall Nano-Hyperspec sensor. (b) DJI Phantom 4 sensor, which collects RGB imagery.

on-board CMOS camera with 2.8-mm optical lens and 12.4 MP resolution was used to collect RGB data [see Fig. 2(b)].

C. Flight Planning

A flight plan was designed prior to each field campaign, considering flight's height, spatial resolution requirements, area to cover, and lighting conditions (see Fig. 3). The Universal Ground Control Station (SPH Engineering, Riga, Latvia) desktop application was used for M600 flight planning and for gimbal control. Regarding M600 flights, hyperspectral swaths were collected with 40% of side overlapping at a speed of 5 m/s and three different heights: 50, 75, and 100 m, originating GSD of 3.1, 4.6, and 6.2 cm, respectively. A total of 15, 8, and 6 hyperspectral strips were acquired, respectively (see Fig. 3). The DroneDeploy (DroneDeploy, San Francisco, CA, USA) was used for Phantom 4 flight planning and its execution. The Phantom 4 was used to acquire 324 RGB images, at 3.4 cm spatial resolution, with 90% along track overlap and 75% lateral overlap, flying at maximum velocity of 6 m/s, and at a height of 80 m from its take-off position.

D. GCP Collection and RGB Imagery Orthorectification

Since the RGB orthophoto mosaic will be used as a reference to align hyperspectral swaths and create the hyperspectral mosaics, GCP have to be used [27], [28]. The UAVs' GNSS receiver does not provide enough positional accuracy (only a few meters). However, not only a good GCP coverage should be ensured but also independent CP should be used to verify the quality of the extracted products [28], [29]. GCPs and CP were measured using a GNSS receiver, in real-time kinematic (RTK) mode, based on the TM06/ETRS89 coordinate system. The imagery geocoding was performed using five GCPs [artificial marks, consisting of 0.5 m diameter circular targets, Fig. 1(c)]—four placed close to each corner and one in the near center of the study area—and the alignment quality assessment was carried out using ten CP (natural marks). The distribution of the GCPs and CP is presented in Fig. 1(b). For this case study, horizontal CP are the most important since they allow to control the planimetric alignment of RGB and hyperspectral outcomes. For a CP of coordinates (East, North), the residuals are calculated by subtracting the coordinates measured by GNSS and the corresponding point interpolated over the RGB orthophoto mosaic. The overall accuracy is given by the RMSE, for n observed CP, as in (1). The mean and the standard deviation can also be determined to assess if some systematic trend may exist in the data

$$\mathrm{RMSE}_{E,N} = \sqrt{\sum_{i=1}^{n} \frac{\left(E_{i,\mathrm{ref}} - E_{i,\mathrm{GNSS}}\right)^2 + \left(N_{i,\mathrm{ref}} - N_{i,\mathrm{GNSS}}\right)^2}{n}}. \tag{1}$$

The RGB data were processed in Pix4Dmapper Pro software (Pix4D SA, Lausanne, Switzerland) to produce a high-resolution orthophoto mosaic. The digital photogrammetric routine implemented in this software includes the following three main stages.

1) A bundle adjustment is performed based on the images' geographic coordinates, obtained from the UAV GNSS receiver, and a set of matching points generated between overlapping images, internal and external camera orientations are also estimated, positional adjustment is performed at this stage using the GCPs.
2) Generation of a dense point cloud (in this case, point density was set to high), based on the camera positions and on the tie points previously computed.
3) Computation of the orthophoto mosaic.

The RGB orthophoto mosaic has a GSD similar to the hyperspectral better resolution (∼3 cm for 50 m flight height). For that reason, the RGB orthophoto mosaic was resampled to the same hyperspectral pixel size (4.6 and 6.2 cm, respectively, for 75 and 100 m flight height) by applying a bilinear interpolation, where the output pixel value is estimated by averaging the surrounding pixels.

E. Preprocessing of Raw Hyperspectral Data

The method presented in this study is applied after a preprocessing of the raw hyperspectral swaths that are orthorectified by Headwall's SpectralView software. The main goal of this step is to correct nonsystematic distortions caused by the turbulence that affects the platform and can cause errors of scale and position. Every rotation around the three main axes (vertical—yaw, transverse—pitch, and longitudinal—roll) is automatically recorded using the on-board IMU. The parametric model used

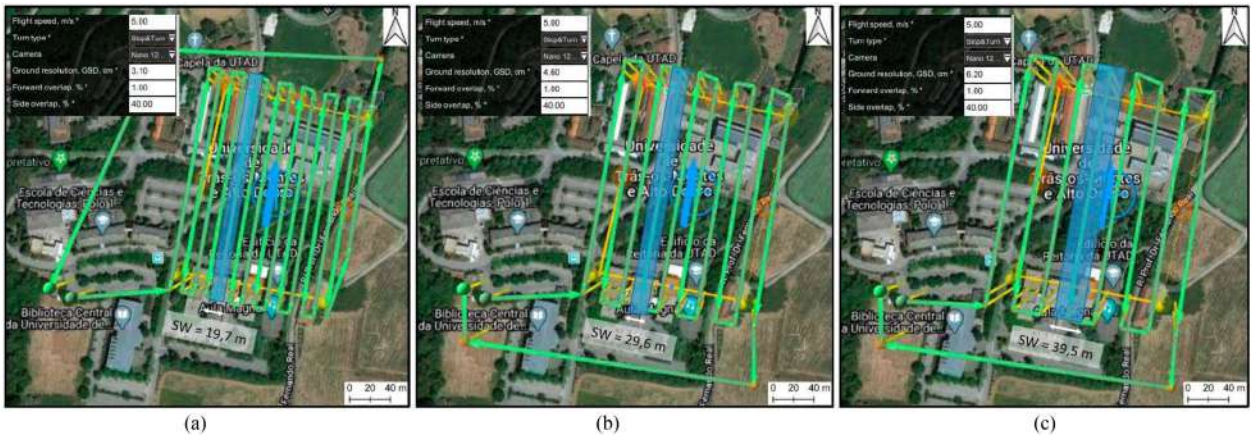


Fig. 3. Hyperspectral flight plans, parameters, and mission area. (a) 50 m height, originating 15 strips with a swath width of 19.7 m. (b) 75 m height, originating 8 strips with a swath width of 29.6 m. (c) 100 m height, originating 6 strips with a swath width of 39.5 m.

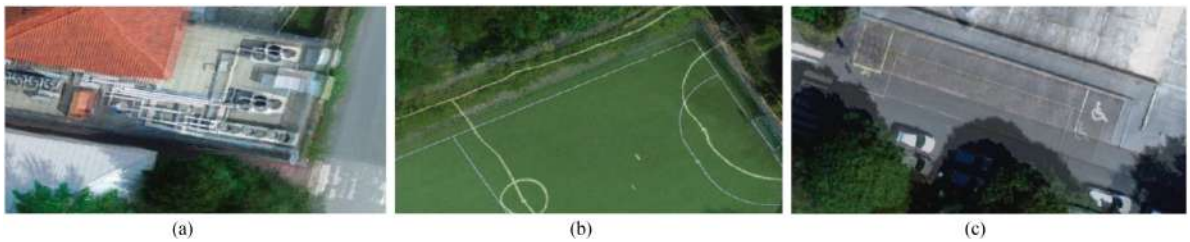


Fig. 4. Some examples of geometric distortions and location errors originated by the preprocessing using Headwall's SpectralView software. (a) Scale and translation errors noticeable in the position of the car and the small grass patch. (b) Translation error. (c) Translation error and distortions noticeable in the lines of the football field.

by SpectralView software (or equivalent) considers GNSS coordinates, timestamps, IMU offsets, the FoV, lens parameters, and sensor orientation, to reconstruct the scanning geometry line by line and to compose each individual swath. The geometric precision of this reconstruction is limited by the GNSS/IMU accuracy leading to strong geometric distortions. Additional iterative processing in Headwall's SpectralView software is required to set the several parameters that better fit the model. The altitude offset parameter is of crucial importance since the relief influences the orthorectification process. Areas with a low absolute altitude and no significant altitude changes allow to set the parameters for multiple cubes simultaneously, which can save a significant amount of processing time. Thus, the use of this kind of approach results in strong geometric distortions, with the generated swaths presenting significant geometric and location errors. Preventing, this way, the use of hyperspectral data in applications that require a precise georectification (e.g., time-series analysis or multisensor data fusion). These issues are clearly visible in Fig. 4, which shows different overlaps between the high-resolution orthophoto mosaic and the resulting orthorectified hyperspectral swaths.

III. Proposed Method

Obtaining push-broom hyperspectral orthomosaic is generally a complex process and involves time-consuming rectification efforts in postprocessing. A significant part of the effort is related to the difficulty in obtaining high geometrically accurate orthomosaics, optimizing the field data collection effort, namely in the acquisition of GCPs. To address this problem, we developed a new approach that simplifies the whole process of obtaining hyperspectral orthomosaics, significantly reducing the time and the necessary computational resources. The method is based on the alignment (registration) between RGB and hyperspectral mosaics from push-broom sensors, which is a very complex and/or time/resource consuming task. In this work, we start from some control points, given by an operator for assessment accuracy purposes, to propose a novel and automatic methodology to: 1) align individual and preprocessed hyperspectral swaths with an RGB orthophoto mosaic; and 2) generate a geometrically accurate hyperspectral mosaic. The developed method is iterative and fully automatic being based on a fast-forward detection of key features in the RGB imagery. Multiple preprocessed hyperspectral swaths to be registered and a high-resolution RGB orthophoto mosaic are the input of the method. From the 270 different bands, only three bands were selected from the visible range (red—670.19 nm, green—540.61 nm, and blue—480.29 nm) are required to be rasterized to generate an RGB swath. This procedure drastically reduces the data to be processed, from several Gb to few Mb per swath. At the end of the process, a transformation model is retrieved, allowing to geometrically transform and merge the original full spectral resolution of the hyperspectral swaths into a continuous hyperspectral mosaic. As presented in Fig. 5, the main steps of

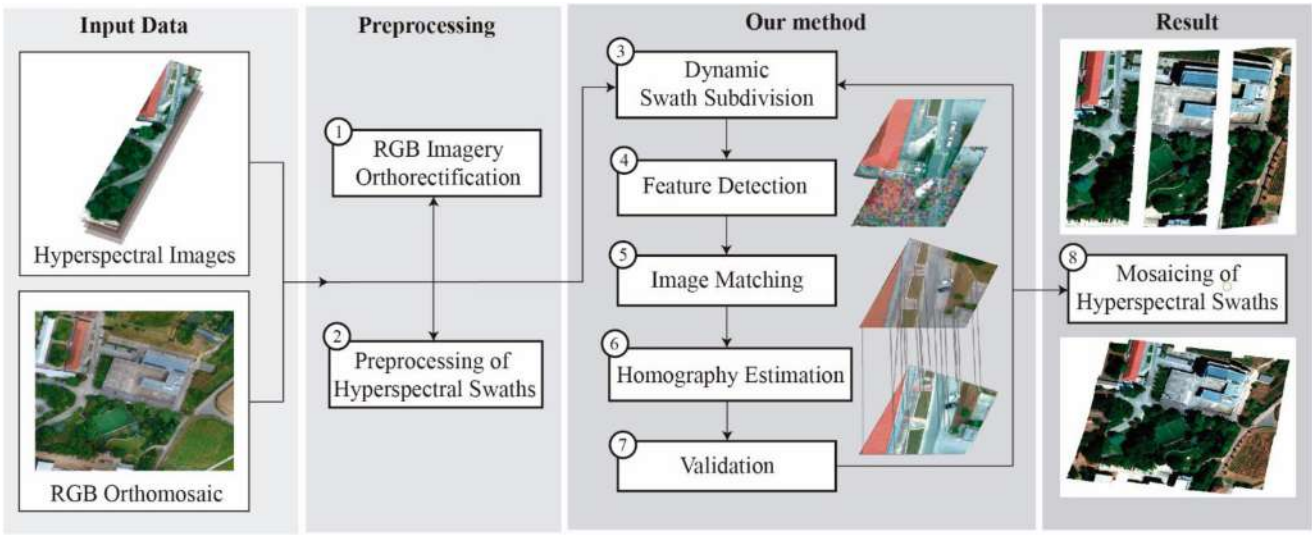


Fig. 5. Workflow of the proposed methodology for the generation of hyperspectral orthomosaics from the alignment of individual and preprocessed hyperspectral swaths.

the method are: 1) image subdivision, 2) feature detection, 3) matching and homography calculation, 4) image transformation, and 5) validation. These processes are iterated to ensure a high accuracy of the resulting hyperspectral orthomosaic. A low initial parameterization is required to properly set up the method's input variables. The development of this solution was carried out using C++ and the OpenCV library [30], and considers as primary goal the efficiency for the ubiquitous execution of the method on portable devices. Some similar methods for image registration can be found in [23] and [31].

A. *Subdivision of Hyperspectral Images*

Geometric deformations are not constant over the swaths, being higher mainly at the edges of the image. For that reason, the first step focuses on dividing the hyperspectral swath into multiple fragments, which equalizes deformations over the fragment. Thus, the geometric transformation will be reached in a fast and accurate way, assuring a plenty alignment with the RGB orthophoto mosaic. Moreover, to ensure an optimal image matching the RGB orthophoto mosaic is also cropped at the same geographical extent, selecting only the area covered by every fragment of the hyperspectral swath. This subdivision is just an intermediate step that aims to simplify the alignment of the images since in each small fragment of area it can be assumed that the deformations are linear. At the end of the process, all fragments are stitched together. This area is set approximately using georeferencing data from both images. Every image fragment must ensure the following three constraints.

1) It should contain more than 50 key-points.
2) The error alignment must be lower than five pixels.
3) At least three control points of the validation mask must be visible.

Regarding these constraints, the length for each fragment must be defined. Initially, this is determined by the HFoV of the hyperspectral camera. Then, the fragment size is adapted in order

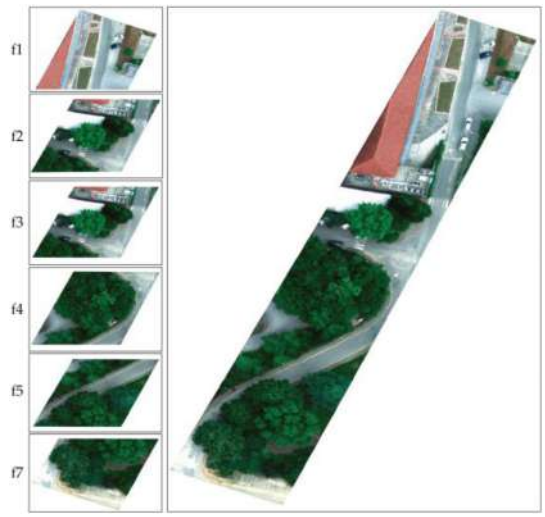


Fig. 6. Example of the application of the first step of the proposed methodology, dividing the original hyperspectral swath into multiple fragments (f#).

to search a high number of key-points and an optimal alignment with the RGB-orthomosaic. This process iteratively seeks to estimate the most adequate image transformation reducing the error resulting from the image registration. This error is obtained by computing the distance between predefined control points of the validation mask, which is set in both RGB-orthomosaic and hyperspectral images. The calculation of this error is described in Section III-E.

Fig. 6 depicts an example of the splitting process of a hyperspectral swath into n fragments (seven in that specific case).

An overlap between consecutive fragments (20% by default) is defined to ensure continuity and to avoid holes close to the fragment's borders. If a fragment cannot be correctly registered on the RGB-orthomosaic geometry, so previous mentioned constraints are not met, the area is extended by 20%, but this value can be customized. This operation is set up to be performed up to

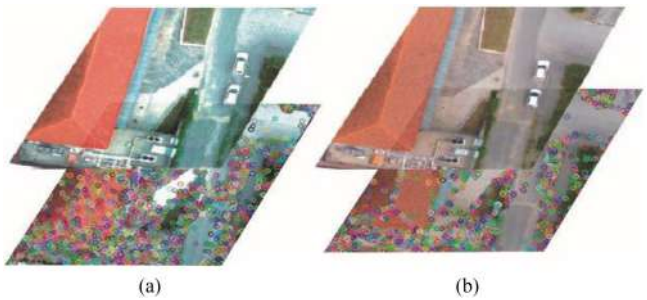


Fig. 7. Detection of key-points in a generated swath fragment. (a) Fragment of the RGB orthophoto mosaic. (b) Fragment of a hyperspectral swath.

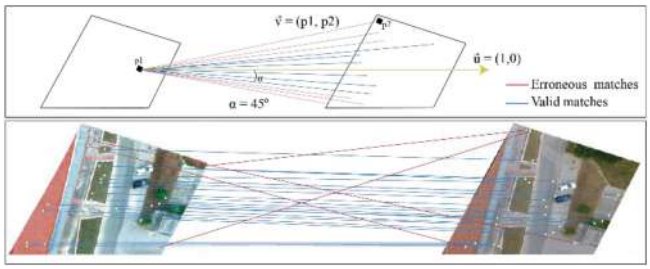


Fig. 8. Key-points matching between a fragment of the hyperspectral swath and the RGB orthophoto mosaic.

five times. More iterations would result in a large fragment and the consequent geometric transformation would be less accurate.

B. Feature Detection

After splitting the input RGB swaths, for each fragment a key-point search process is carried out, using the Oriented FAST and Rotated BRIEF (ORB) [21] method. Unlike SIFT [32] ORB is relatively immune to Gaussian image noise. Furthermore, ORB is an efficient method and is not demanding in terms of hardware. This method was compared to SIFT and SURF in terms of the performance. The overall time executing of the program decreases $\times 10$ for launching SIFT and SURF algorithms [33]. Considering the key-points detection, ORB also presents a more accurate recognition since most points are considered valid for the matching process (90%). According to the results of this experiment, ORB was selected for the image feature detection. This step is influenced by the characterization of the study area. In this work, the region covered contains buildings, roads, dense vegetation, vehicles, and other human-made objects. The detection of key features is one of the most relevant stages to ensure an optimal image registration. The goal is to achieve a balance between the number of key-points and efficiency. In contrast to heterogeneous regions, where geometrical shapes or significant color changes are presented, in those image fragments that present a homogeneous appearance, a greater effort is required to find optimal key-points. After multiple tests, to ensure the recognition of key-points in both scenarios, the developed method was set to detect 10 000 points per fragment.

As a result of this step, a set of ORB key-points is collected for each fragment (see Fig. 7), which will be analyzed in the next step to match points that represent the same feature in both images and then allowing the calculation of the geometric transformation.

C. Image Matching and Homography Calculation

In this section, the key-point matching process between the hyperspectral swaths and the RGB orthophoto mosaic is described. The key-points found in a given fragment will be used to estimate the transformation matrix to be applied per pixel, calculating the homography. Thus, the main objective is to search matches in both images, between the detected key-points. For this purpose, the k-nearest neighbors (KNN) algorithm is applied to search for the *k* best matches based on the calculation of the Hamming distance [34], between pairs of key-points. In this specific case, the KNN method is parameterized to provide the two best matches for each key-point. The first match is considered valid if its distance is 10% lower than the second one. If this constraint is not satisfied, the match is discarded. As a result, a list of matches is provided but these are not definitive. A new filter, based on the inclination of vector (v) formed by the first (k1) and second (k2) key-points, is proposed in this method. Due to the fact that both fragments are oriented to the north and represent the same region of the study area, most matches should present almost horizontal inclination. To avoid erroneous matches, the angle (α) between **v** and **u** (1,0) is calculated. Then, if the angle is greater than 45°, the match is discarded. This threshold has been selected after multiple tests and considering the deformations of the hyperspectral swaths. Fig. 8 shows an example of matches that were discarded and accepted for the application of the method.

Once the matches are obtained, the next step consists of the calculation of the homography. The homography is represented by a 3×3 matrix and is widely used for different purposes such as the estimation of the camera pose from coplanar points, perspective removal, and image registration. In this study, the homography is calculated to fit the position of every pixel of the hyperspectral fragment allowing to achieve a better alignment with the RGB orthophoto mosaic used as reference. However, in this step, the second constraint mentioned before must be satisfied. To consider a valid homography, the resulting error in the alignment must be lower than five pixels. The error is calculated using the control points of the validation mask. This test is performed after the homography is calculated. This validation process is described in more detail in Section III-E. After homography validation, a geometric transformation (translation, rotation, and scale) is applied for each pixel in the fragment.

D. Special Cases

The method considers that some regions of the hyperspectral swaths could present a higher difficulty to be aligned with the RGB orthophoto mosaic due to the presence of dense vegetation, homogeneous appearance of objects or some irregular reflections. Following those special cases, the partition of images is not static and it is modified depending on the area observed for each fragment. If no key-points are detected or if the error in the alignment is higher than the defined threshold (five pixels), or no control points are visible, the length of the fragment is

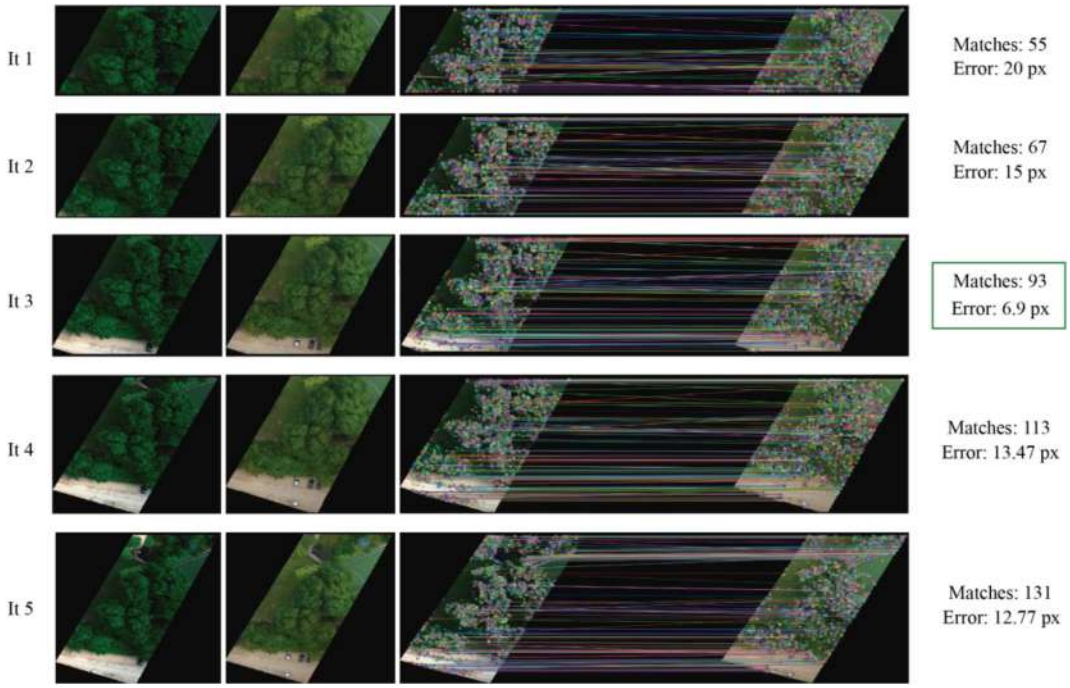


Fig. 9. Automatic fragment resizing to estimate the most accurate image registration in regions with homogeneous appearance.

automatically resized. Fig. 9 shows an example of the surveyed area where a homogeneous region with vegetation coexists. The process to resize the fragment can be iterated until five times. The first step is to extend the cover area downwards in order to detect a higher number of valid matches and to reduce the error alignment. If it is not possible to cover a larger area downwards the method seeks to look for new points in previously aligned areas. The stop criterion is defined when reaching an error alignment lower than five pixels. If it is not satisfied, the iteration in which the lowest error is achieved is selected to estimate the homography from the detected matches. In this case, the method continues by the end of the fragment corresponding to the chosen iteration. The consideration of these special cases supposes a significant improvement of the quality of the orthomosaic regarding the complexity of such regions that do not contain enough changes and key features.

E. Validation

To validate the alignment between the RGB orthophoto mosaic and hyperspectral swaths, a validation mask for each image is created where control points are represented as unique color combinations. Each control point is located in meaningful places such as corners, and key-objects that can easily be recognized in both hyperspectral and RGB imagery. The distribution of these markers is carried out by randomly considering the following constraints: at least 20 control points have to be included per swath, and distributed over the image to ensure that each fragment can be validated and these cannot be collinear to enable an automatic recognition. At least, two or three control points are detected for each fragment. These control points are not used for the alignment process since they are not enough to estimate the homography parameters, namely, more image features must be detected in the fragment. The goal is to validate the accuracy of the resulting image registration. Fig. 10 shows an example of markers used to validate the geometric transformation for each fragment of hyperspectral images.

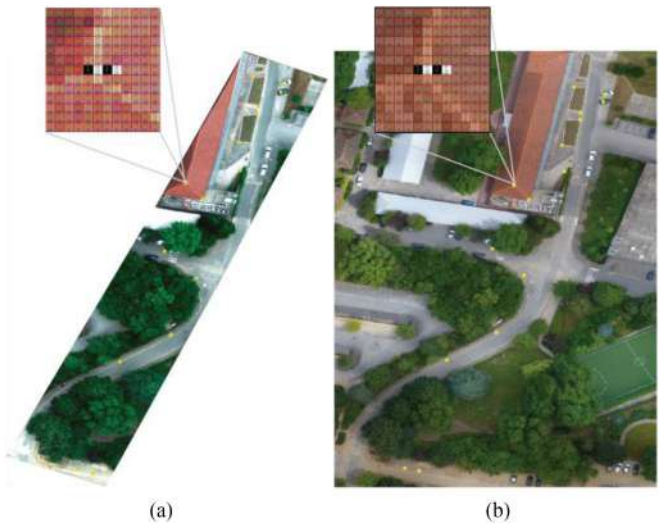


Fig. 10. Distribution of the manual control points used to compute the alignment accuracy. (a) Hyperspectral swath. (b) RGB orthophoto mosaic.

Once the validation masks have been created, the method is capable of registering the visible markers for each fragment. Then, the resulting image transformation from the estimated homography is applied and the Euclidean distance is calculated from the position of the projected marker to its corresponding pixel in the RGB orthophoto mosaic. Fig. 11 shows the

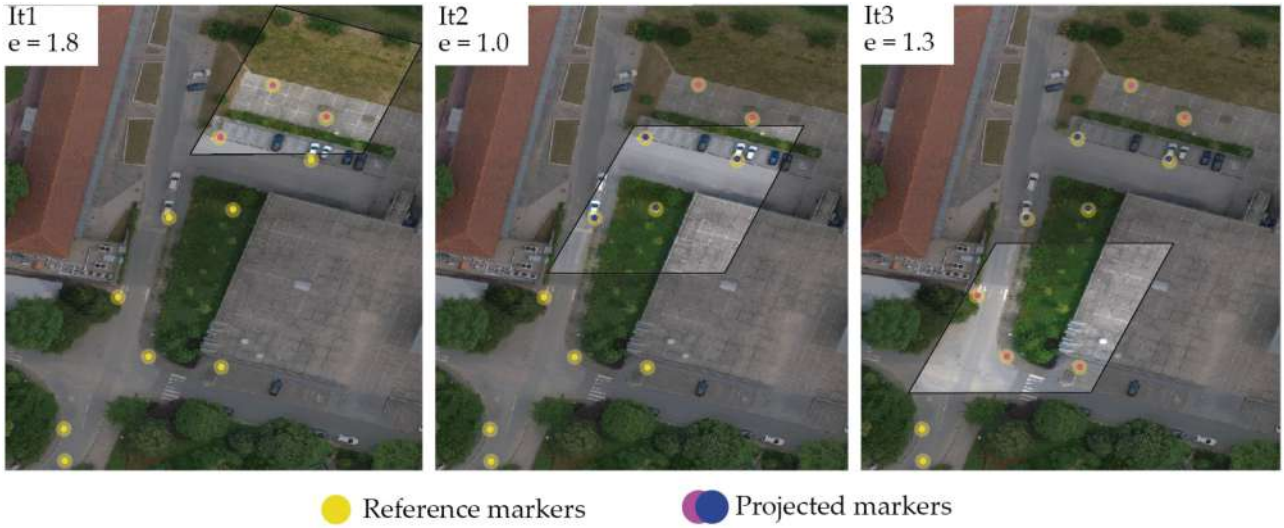


Fig. 11. Markers detection from validation masks and reprojected markers for each hyperspectral fragment. The parameter *e* represents the alignment error in pixels.

projection of visible markers for a set of fragments from a hyperspectral swath on the corresponding markers in the RGB orthophoto mosaic. The distance (d) between two markers is used to calculate the error of image registration.

IV. Results

A. RGB Orthorectification Accuracy

The RGB orthophoto mosaic that serves as a reference for the coregistration of the hyperspectral swaths was processed using Pix4Dmapper Pro software and integrating GCPs. The accuracy quality checking of the performed orthorectification was carried out by using the ten CP distributed in the area as presented in Fig. 1. Analyzing the values that were obtained for the mean residuals—the mean value is close to zero—it can be concluded that there are no systematic errors. The results presented in Table I also allow us to conclude that the geometric adjustment reaches a subpixel accuracy (∼0.3 px), corresponding to about 1 cm RMSE. From the qualitative point of view, a visual inspection leads us to conclude that sizes and shapes of objects are well preserved. Linear features like roads and buildings are continuous and well defined.

TABLE I
Geometric Quality of the RGB Orthophoto Mosaic Used as Reference for Hyperspectral Swaths Coregistration

Check point	$\Delta E\ (px)$	$\Delta N\ (px)$	ΔEN	
			(px)	(cm)
1	-0.376	0.780	0.866	2.94
2	0.456	-0.863	0.976	3.31
3	0.181	-0.464	0.498	1.69
4	0.038	0.142	0.147	0.49
5	-0.631	0.433	0.765	2.60
6	-0.467	-0.872	0.989	3.36
7	0.502	0.461	0.681	2.31
8	-0.518	-0.095	0.526	1.79
9	0.668	0.641	0.926	3.14
10	0.360	-0.332	0.490	1.66
Mean residuals	0.021	-0.017	-	-
RMSE	0.457	0.572	0.257	0.87

B. Performance of the Automated Coregistration

The different orthomosaics from the study area have been created using a laptop with the following hardware specification: Intel i7-8565U ×64 CPU, 16GB RAM, and GeForce GTX 1050 GPU. The proposed method has been tested considering three flights at different heights: 50, 75, and 100 m. For each flight, the number of captured hyperspectral images and the pixel size on the ground (GSD) change (see Fig. 3). Thus, the method's performance is assessed with three sets of hyperspectral images, allowing to estimate the impact of flight height (thus, GSD) on the overall accuracy. Fig. 12 shows an overall overview of the results obtained from the proposed method at the different flight heights that are presented in the following sections. Moreover, the results of corrected images from some regions of the study area and the RGB orthomosaic are shown in Fig. 13.

1) Flight at 50 m: According to the first flight, 15 hyperspectral swaths (∼20 m width) were processed to generate

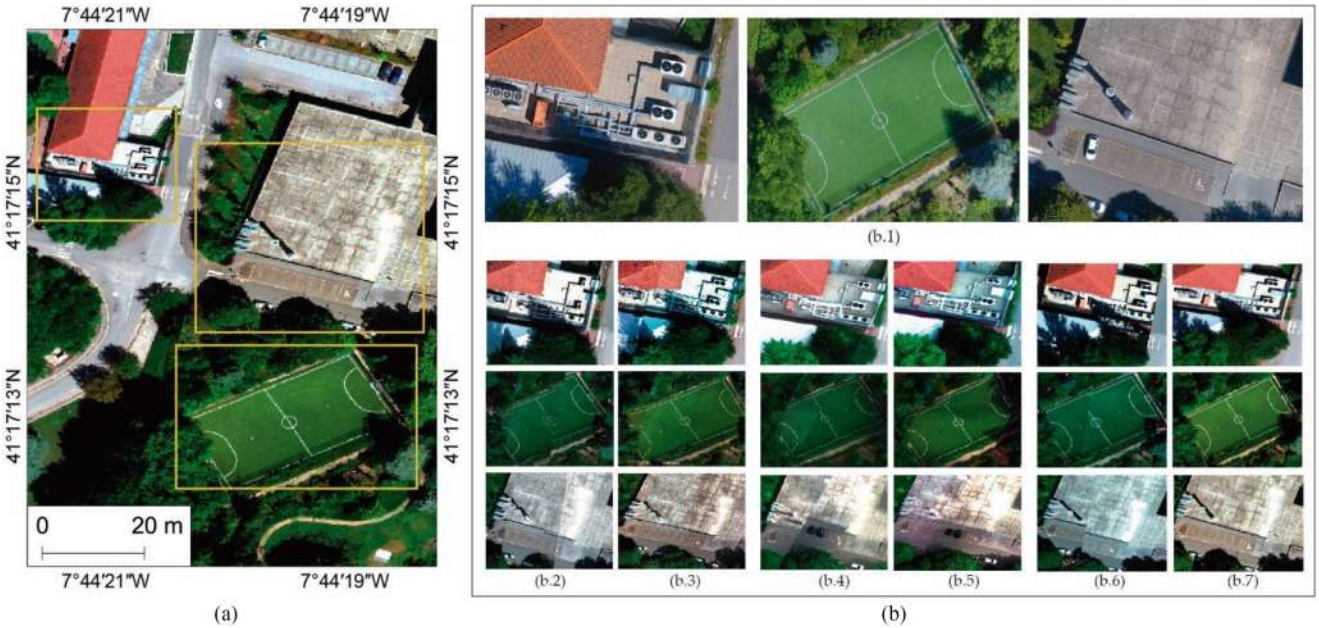


Fig. 12. Overview of the results from the (a) flight at 100 m height and (b) close-up views of the polygons highlighted. (b.1) RGB orthophoto mosaic. Hyperspectral orthomosaics before applying the method for the flights at (b.2) 50 m, (b.4) 75 m, and (b.6) 100 m. Hyperspectral orthomosaics after applying the method for the flights at (b.3) 50 m, (b.5) 75 m, and (b.7) 100 m.

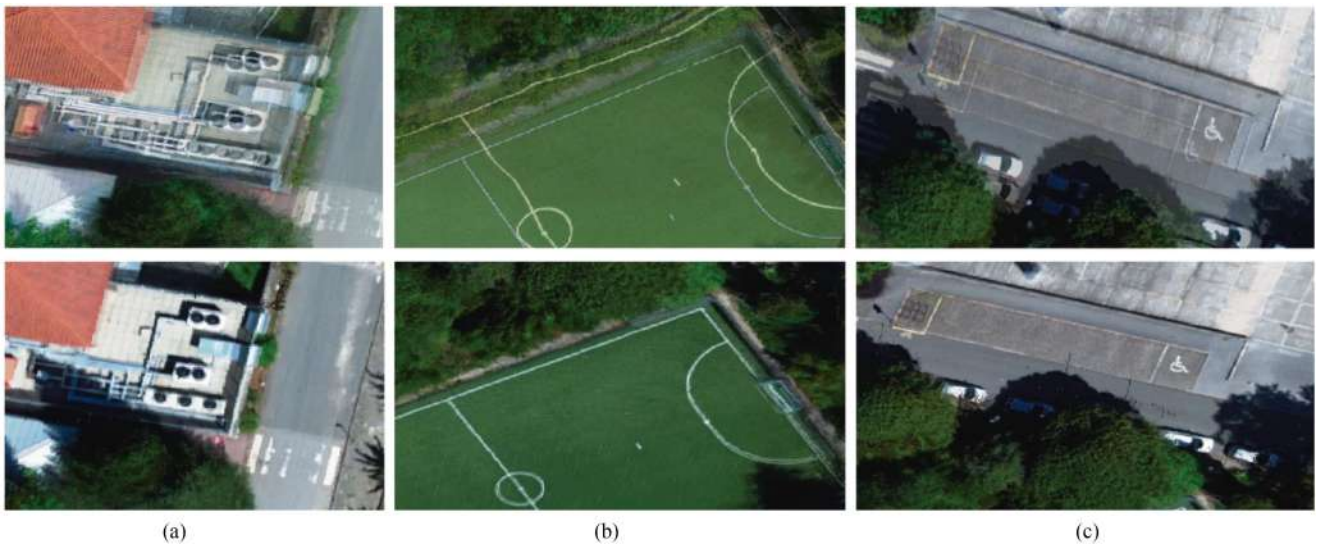


Fig. 13. Resulting hyperspectral image alignment to the RGB orthomosaic of three different regions in the study area. The top row represents the overlap of the input hyperspectral images and the bottom row represents the images corrected by the proposed method.

the orthomosaic. Table II presents the results obtained by applying the proposed method in terms of the accuracy of the image correction. The relative positional accuracy between each georectified hyperspectral swath and the correspondent RGB orthophoto mosaic is determined by calculating the RMSE, determined by calculating the Euclidean distance between the rectified coordinates in the hyperspectral swath and the reference coordinates in the RGB mosaic (validation marks). From Table I, it can be concluded that accuracy increases with the number of fragments. In fact, to keep a low error there are advantages of using smaller fragments. The variation in the number of fragments per hyperspectral swath is variable, as our method adapts to the characteristics of the area, starting with a value depending on the swath's width, the overall error affecting this orthomosaic is of 4.63 px (∼14 cm).

Fig. 12(b.2) shows the hyperspectral mosaic computed by overlapping all hyperspectral swaths after the preprocessing carried out in SpectralView software and Fig. 12(b.3) presents the hyperspectral data after corrections applied from the proposed method. Focusing on the visualization of borders of the football field, roads, and roofs, a high alignment is achieved with the proposed method. Regarding these results, the method works properly but a higher number of iterations is required to determine the size of the fragment, which is usually extended

TABLE II
ACCURACY ASSESSMENT FOR THE 50 M FLIGHT HEIGHT HYPERSPECTRAL SWATHS, CONSIDERING THE MEAN ERROR OF EVERY IMAGE CORRECTION AND FRAGMENTS REQUIRED TO BE ALIGNED

Hyperspectral swath	Mean error (px)	No. of fragments	F1	F2	F3	F4	F5	F6	F7	F8
1	7.79	5	1.42	4.44	1.85	18.50	12.74	-	-	-
2	5.62	5	0.90	1.67	2.43	18.29	4.82	-	-	-
3	4.48	4	4.8	4.1	4.50	4.38	-	-	-	-
4	2.70	7	1.07	0.62	3.11	2.97	3.81	4.44	2.87	-
5	2.66	5	2.58	1.30	2.69	0.97	5.78	-	-	-
6	2.72	6	0.88	0.84	1.74	3.06	1.63	8.20	-	-
7	8.68	5	1.70	4.99	0.68	27.25	8.76	-	-	-
8	1.92	6	0.90	0.71	0.59	4.75	3.01	1.58	-	-
9	2.13	7	3.61	1.03	1.2	0.21	3.94	2.23	2.68	-
10	1.86	8	2.74	4.13	0.32	1.26	1.19	1.73	0.29	3.20
11	3.95	5	1.48	4.79	0.69	7.57	5.24	-	-	-
12	10.64	3	1.61	10.3	19.9	-	-	-	-	-
13	3.00	5	1.14	4.93	1.11	3.97	3.83	-	-	-
14	9.97	6	0.89	7.99	4.91	11.24	11.57	-	-	-
15	1.45	8	0.80	0.97	3.2	3.9	5.43	6.50	1.06	1.45
Total	4.63	85								

to achieve a better adjustment. This implies a negative impact to the alignment as shown for the seventh and tenth swaths (see Table II) and also an increase of the computational effort, as shown in Section IV-E.

C. Flight at 75 m

In this test, the study area is covered at a higher flight (75 m, 4.6 cm GSD), keeping the same acquisition plan. Consequently, only eight hyperspectral swaths were acquired (width ∼30 m). Table III presents the alignment errors and number of fragments used to process every hyperspectral swath. Most swaths were divided into six or seven fragments, resulting in a more homogeneous distribution of individual error, ranging from 2.37 to 5.00 px. The mean error of the resulting orthomosaic is 3.82 px, which corresponds to ∼17 cm. At a higher flight height, the method is able to find more valid matches and key-points to obtain the homography. As a result, pixels of hyperspectral images are correctly transformed and the image deformation is plenty removed.

Fig. 12(b.5) shows parts of the resulting mosaic (see Fig. 12(b.1) for the RGB orthophoto mosaic of the same areas and Fig. 12(b.4) for the hyperspectral data before alignment). In contrast to the 50 m height hyperspectral mosaic, this one presents a higher accuracy, in terms of pixels, but some challenging parts where homogeneous surfaces appear and the alignment is a bit worse remain. This is discussed in more detail in Section V.

D. Flight at 100 m

In this third flight, a hyperspectral dataset composed of six hyperspectral swaths (width ∼40 m) acquired at 100 m flight height (6.2 cm GSD) was aligned. Considering a larger area covered for each image, the method automatically focused on reducing the size of the fragment to better fit the geometric transformation. This way, most swaths are divided into nine fragments and even two of them in ten. Table IV presents the results for this coregistration process. Regarding the mean error, it is 2.99 px (∼18 cm).

TABLE III
ACCURACY ASSESSMENT FOR THE 75 M FLIGHT HEIGHT HYPERSPECTRAL SWATHS, CONSIDERING THE MEAN ERROR OF EVERY IMAGE CORRECTION AND FRAGMENTS REQUIRED TO BE ALIGNED

Hyperspectral swath	Mean error (px)	No. of fragments	F1	F2	F3	F4	F5	F6	F7	F8
1	2.54	6	1.56	2.54	3.82	2.02	2.85	2.54	-	-
2	3.03	8	1.86	1.64	3.22	1.06	3.55	3.11	4.83	5.00
3	4.79	7	1.29	4.06	1.13	4.51	3.06	6.9	4.79	-
4	4.12	6	5.54	3.04	3.84	4.28	3.08	4.9	-	-
5	2.37	7	2.92	0.96	2.62	1.75	2.50	3.83	2.03	-
6	4.13	5	4.8	4.05	1.39	1.82	8.5	-	-	-
7	4.60	7	4.7	2.38	1.30	3.5	1.97	10.65	7.9	-
8	5.00	5	7.51	2.75	3.98	1.44	9.42	-	-	-
Total	3.82	51								

TABLE IV
ACCURACY ASSESSMENT FOR THE 100 M FLIGHT HEIGHT HYPERSPECTRAL SWATHS, CONSIDERING THE MEAN ERROR OF EVERY IMAGE CORRECTION AND FRAGMENTS REQUIRED TO BE ALIGNED

Hyperspectral image	Mean error (px)	No. of fragments	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	2.54	9	3.39	1.51	4.77	3.84	6.06	3.58	1.13	3.98	3.40	-
2	2.38	10	0.77	1.57	4.81	3.04	1.55	0.88	4.88	2.32	1.82	2.19
3	1.59	10	0.75	0.34	1.41	0.78	1.53	2.26	0.57	3.73	3.16	1.36
4	3.00	9	2.32	4.04	4.94	1.32	2.57	1.36	1.00	2.85	6.6	-
5	4.57	8	1.08	16.9	2.31	2.75	3.04	3.20	2.44	4.77	x	-
6	2.30	8	1.33	0.92	10.4	3.22	0.99	2.13	3.82	1.72	2.30	-
Total	2.99	54										

Fig. 12(a) shows part of the hyperspectral orthomosaic [detailed views at Fig. 12(b.7)]. At this scale, an almost plenty alignment with the RGB orthophoto mosaic used as reference [Fig. 12(b.1)] is visible. In addition, those challenging parts could be corrected by using surrounding key-points that enable the extraction of valid matches and thus, the calculation of an accurate image transformation. The mean error per fragment ranges between 1.59 and 4.57 px.

E. Performance Analysis

The proposed approach is able to achieve spatial accuracies that compare well with those obtained identifying matching points manually. Thus, the processing time is one of the main factors to be considered for choosing an automatic approach. For that reason, the three hyperspectral datasets were fully processed using Headwall's SpectralView software, and the time was noted. This tool provides a semiautomatic method since manual parameters need to be adjusted after each iteration. Table V compares the timing measurements per dataset and method used. The manual identification of validation markers to check the quality alignment in the proposed method was also considered, with an average time of 3 min. Comparing both methods, the new approach requires only a few minutes against several laborious hours of work, representing a saving in time of 97%, 99%, and 99%, respectively, for flight heights of 50, 75, and 100 m. The

TABLE V
COMPARISON OF EXECUTION TIMES TO GENERATE EVERY ORTHOMOSAIC USING THE PROPOSED METHOD AND HEADWALL'S SPECTRALVIEW SOFTWARE

Orthomosaic	Proposed method			SpectralView (hours)
	Markers positioning (min)	Processing time (s)	Total (s)	
Flight at 50 m		1141.64	1321.64	10.5
Flight at 75 m	3	57.09	237.09	7
Flight at 100m		78.81	258.81	6

automated approach presented in this study represents for this specific case study a saving of 23 h of work, requiring only 30 min to process the three hyperspectral mosaics. Regarding the individual automated processing, a greater time (∼20 min) was necessary to generate the first mosaic, which is formed by aerial images at a lower flight height. This mosaic required a higher number of iterations to fit the size for each image fragment. This is due to the fact that a smaller region is covered by the sensor (swath width) and there are fewer key-points detected and less good matches can be found. The fastest execution is obtained for the second mosaic (less than 1 min if excluding manual identification of validation marks) since no more than seven fragments are required for most hyperspectral images. A good balance is achieved considering the observed area and level of details. Finally, the generation of the third orthomosaic is a bit slower (∼1 min 20 s), since the length for each fragment is shorter to fit the most accurate alignment, namely, the number of fragments for each image increases.

As a summary, the most relevant factor that influences the method's performance is the number of iterations required to fit the size of the fragment. Considering this variable, the proposed method works faster for flight heights over 75 m, corresponding to swath widths of 30 m and a 4.6 cm GSD. At this flight height, the covered area is large enough to achieve a precise orthorectification by the alignment with the photogrammetric orthophoto mosaic used as reference. In terms of effectiveness, these results compare well with those obtained by Rublee *et al.* [21]. However, the method proposed in this article is much less demanding in terms of resources and time.

V. DISCUSSION

A. Hyperspectral Mosaic's Reconstruction

In recent years, the proliferation of hyperspectral data acquired by push-broom sensors carried by UAVs has led to the development of several semiautomated [5], [6], [11], [13] and automated [17], [18] processing methods. However, still there are some challenges remaining related to field operations and data processing (data acquisition, quality assessment, algorithms' optimization, etc.). Associated with these factors, it is still necessary to consider the large volumes of data originated by hyperspectral sensors. All of these factors combined make it necessary to develop precise, expedited, and automated methods to generate hyperspectral mosaics from UAV-based push-broom swaths. In this study, we present an innovative automated methodology able to address these challenges, speeding up the coregistration strategy based on the establishment of direct transformation between hyperspectral swaths and an RGB orthophoto used as reference. In addition, the proposed approach significantly reduces field work since it only requires a limited number of GCPs to georectify the RGB mosaic. From the theoretical point of view, the method maintains its performance even without using any GCP, since it is based on the identification of key-points, regardless of the precision of the reference orthophoto mosaic. In fact, some UAV manufacturers started to offer solutions with embedded RTK at a reasonable price, which can turn the use of GCPs not necessary.

Comparing to the actual state of the art, most methods that use RGB orthophoto mosaics as a reference to georectify hyperspectral data, Habib *et al.* [18] used the same hyperspectral sensor employed in this study (17-mm lens) to acquire 5 cm GSD swaths. The authors used the high-resolution RGB DEM to partially correct from geometric deformations and then applied a modified version of the SURF algorithm for coregistration. The final results achieved an overall accuracy ranging from 0.5 to 0.9 m RMSE per swath. Considering the three datasets used in this study, with comparable resolutions, it was possible to achieve accuracies in the order of 0.15 m. This substantial improvement in accuracy relies on the division of swaths into fragments, depending on the various parameters of the image and the study area. Another decisive factor is that we used an RGB combination for the hyperspectral swaths, whereas most methods use only a single band (usually the red band), which prevents radiometric differences to be considered. The main innovation of the method proposed in this study in comparison with previous works can be summarized as follows.

1) The development of an automatic workflow based on some control points given by an operator.
2) An iterative process to fit the size of image fragment in order to achieve an accurate image registration.
3) Correction of image deformation even in homogeneous areas where, for instance, vegetation and uniform surfaces coexist.

4) The efficiency for image processing and computing of geometric transformations.

Angel *et al.* [22] proposed an innovative method exactly for the same purposes as ours. However, although the excellent results globally provided by this automated method, it presents as major limitations the fact that important resources were required and some hours were needed to conclude the procedure. Globally, our approach is comparable in terms of accuracy but significantly better performance in terms of resources and processing time.

New approaches are based on the use of LiDAR systems to simultaneously acquire hyperspectral and LiDAR point clouds data facilitating data fusion [35], [36]. Moreover, other methods are proposed as a practical tool for RGB UAV in shadowed and saturated zones [37]. These approaches present relevant drawbacks such as a more complex definition of the acquisition plan, limitations of LiDAR resolution, a bigger aerial platform to support the weight of both sensors and a high hardware expense. The results obtained from the application of the proposed method enable us to avoid the use of more expensive alternatives such as LiDAR sensors. In fact, to overcome the challenges from its push-broom sensor, Headwall offers a solution to use both LiDAR and hyperspectral sensors to simultaneously acquire data in a single flight [38]. Several studies apply the same hyperspectral sensor as the one used in this study in combination with a LiDAR sensor for the semantic classification of real-world scenarios, generation of hyperspectral mosaics and 3-D point clouds, and the documentation of complex scenarios with geometric and spectral attributes [9], [39], [40]. However, on top of the costs of the acquisition of a hyperspectral sensor along with a UAV capable of supporting it, this option raises, even more, the acquisition costs of such a solution, increases the UAV payload, reducing significantly the maximum flight time.

Furthermore, solutions based on snapshot sensors to collect hyperspectral data have been proposed [41]–[45]. In contrast to push-broom sensors, snapshots enable the capture of full images for each shot instead of a line-by-line sweep and can be photogrammetrically processed. This option presents an important advance in the field of hyperspectral imaging but for the time being the available sensors provide a lower spectral resolution. Push-broom sensors are able to capture a high number of spectral bands but as a disadvantage, a higher image deformation is presented.

Our solution can be applied for most scenarios in the real world. In terms of the method's parameterization, only two parameters may be modified: the number of key-points to be detected for each fragment ($\sim$10 000), and the maximum angle (45°) to consider a valid match. The values of both parameters have been considered valid for most types of possible scenarios since the surveyed area represents heterogeneous regions characterized by both saturated zones by dense vegetation as well as clear areas with homogeneous surfaces. Regarding the sensitivity of these parameters, the target number of key-points directly influences the detection of valid matches. Generally, a lower value implies a worse registration. However, specific areas with many significant features enable a faster registration considering less key-points. This value could be reduced until 3000 points for these areas. According to the maximum angle to obtain valid matches, it could be ranging from 30° to 60°. A lower value enables matching more distorted key-points. If the distortion of input hyperspectral swaths is too high, lower values are recommended. Other parameters such as the fragment overlap and the fragment size should not be changed to ensure the correct method applicability. For instance, a larger size of the fragment produces a worse registration with the RGB orthomosaic and also this will negatively influence the final performance. Moreover, if the fragment overlap is reduced a worse registration will be presented in the borders of fragments.

B. Accuracy of Orthomosaic

The alignment errors from the results of this study show that the method achieves a high accuracy if the hyperspectral swath can be divided into many fragments. Due to the fact that the geometric deformation is different for each part of the image, the method can estimate in a more precise way the transformation for shorter fragments. The second variable that influences the accuracy of the orthomosaic reconstruction is the flight height. Consequently, some inevitable local distortions are presented in remote sensing images caused by eighter the variation of the ground surface or the changing position of the viewpoint [46], [47]. The proposed method is based on feature extraction for each fragment in order to match pairs of pixels from hyperspectral images and the RGB orthophoto mosaic, which was obtained as mentioned in Section II-D. To ensure the estimation of a precise homography, the image fragment must contain meaningful features that describe the cover area. According to the obtained results, the alignment error is lower at a higher flight height. This correlation is due to the larger area covered in each fragment and even a lower object deformation is observed from a farthest viewpoint. By applying the proposed method, the accuracy of orthomosaic is always under five pixels for all performed flights (50, 75, and 100 m). Considering the relationship between flight height and the efficiency of the method, this last decrease if no key-points are detected and multiple iterations must be launched to fit the size of the image fragment. Regarding our results, the processing of images captured at 75 m represents the most adequate balance in terms of accuracy (3.8 px, $\sim$17 cm) and execution time (57 s).

Other methods explored different approaches with the goal to improve the geometrical accuracy of hyperspectral data acquired from push-broom sensors. Habib *et al.* [48] explored different boresight calibration approaches of a push-broom sensor; the RMSE values decreased from 0.3, 0.6, and 1 m, respectively, at X, Y, and Z coordinates when not applying any corrections to 0.03–0.11 m in X, 0.01 and 0.06 m in Y, and 0.17–0.60 m when using a rigorous approach based in tie-features. Lenz *et al.* [49] proposed a method using SURF for tie-point estimation on push-broom scanners. The method was tested in two datasets with approximately 0.5 m GSD. Residual RMSE was 1.5 m for the first dataset (forest area) and 0.5 m for the second dataset (urban area). In [35], point-pairing approach was used for bundle adjustment of hyperspectral push-broom imagery and the integration with LiDAR data to calibrate the two sensors

was conducted. This approach enables the authors to reduce the RMSE on *X*, *Y*, and *Z* from 0.273, 0.538, and 0.870 m to 0.054, 0.035, and 0.352 m. Turner *et al.* [7] reduced the errors and achieved an error of approximately reduced the error from 0.20 to 0.05 m. In [50], a modified SURF-based feature matching approach was used to align partially rectified hyperspectral mosaics with RGB orthophoto mosaics. With this approach, the authors improved the hyperspectral geometric accuracy of approximately 5–0.6 m. And in [18], the accuracy was improved from 2.5 to 0.25 m using tie points and linear features from RGB orthophoto mosaics.

On the other hand, studies using the traditional preprocessing presented a higher misalignment. Jaud *et al.* [51] evaluated direct georeferencing of hyperspectral data acquired from push-broom sensors in two coastal environments; horizontal RMSE values ranging from 1.4 to 2.3 m were obtained. RMSE errors of approximately 1 m were obtained in [9] when using the conventional data preprocessing pipeline. Arroyo-Mora *et al.* [52] obtained an error of 2–3 geocorrected image pixels in the easting and northing directions, respectively, 0.113 and 0.107 m when evaluating the usage of push-broom sensors for ecological monitoring. Moreover, relevant methods were presented focused on image processing and registration based on establishing robust point correspondences between two sets of points [53] and seeking reliable correspondences between two feature sets [54]. These advances in the field of computer vision demonstrate the capabilities of image feature detection and matching strategies considering also nonrigid transformation for image registration.

The proposed methodology provides an expedite solution able to produce a hyperspectral mosaic with an accuracy on the order of two to five times the ground sampling distance of the high-resolution RGB mosaic used as reference. The complex area used was covered by several individual hyperspectral swaths and was processed in only a few seconds (around 1 min for swaths acquired at 75 and 100 m flight height) using a regular laptop. Thus, the proposed method implements a simple coregistration strategy and achieves high positional accurate results. In fact, the use of an RGB orthophoto mosaic discards the usage of targets in the field for finding key-point for hyperspectral data alignment since those can be easily identified in the RGB imagery.

The method can be applied to align hyperspectral data in accordance with photogrammetric outcomes, bringing the possibility for accurate spectral extraction on a given area. Moreover, it is possible to create 3-D hyperspectral point clouds. Also, multisensor data fusion can take place making it possible to accurately conduct data extraction of biophysical parameters for a plant- or tree-level analysis, to combine geometric, thermal, and spectral data in the same range, and perform time-series analyses. The versatility of the proposed method and its capability to adjust the size of the image fragment to obtain the most adequate balance between the image recognition and the image deformation enable the efficient creation of accurate hyperspectral mosaics based on push-broom sensors.

VI. Conclusion

The method presented in this study proved to be effective for the generation of hyperspectral mosaics using a set swath captured by a push-broom sensor. Even though there are several challenging parts of the images where vegetation and homogeneous areas coexist, the method divides the input image into several fragments to ensure both the recognition of meaningful features on the image and the most accurate correction of image deformation. This iterative approach focuses on key-points detection between hyperspectral swaths and the RGB orthophoto mosaic used as a reference, to compute enough number of valid matches. Then, every image fragment is transformed by applying the resulting homography. Thus, the proposed methodology means a significant advance in the field of hyperspectral remote sensing to overcome the limitation to generate precise hyperspectral mosaics using push-broom sensors. The automatic workflow and the ease to be replicated will lead to the proposed method as a standard that will enhance current research in the field of hyperspectral remote sensing.

The applicability of the proposed method can be extended to other types of purposes related to the multisource data fusion. The use of other remote sensing data such as multispectral or thermal infrared can be interesting for several tasks such as material segmentation, environment understanding, tree species identification, and disease estimation. As further research, nonlinear transformation models will be studied to improve the image registration in challenging scenarios. Moreover, the proposed method enables the development of a promising research line related to the generation of hyperspectral 3-D point clouds in a cost-effective manner when compared to LiDAR point clouds since photogrammetric point clouds can be used. The study of the distribution of hyperspectral data in 3-D environments opens new possibilities to extract meaningful information from real-world scenarios.

References

[1] A. F. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, Jun. 1985.

[2] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs," *Remote Sens.*, vol. 5, no. 2, pp. 949–981, Feb. 2013, doi: 10.3390/rs5020949.

[3] M. F. McCabe *et al.*, "The future of earth observation in hydrology," *Hydrol. Earth Syst. Sci.*, vol. 21, no. 7, pp. 3879–3914, Jul. 2017.

[4] T. Adão *et al.*, "Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry," *Remote Sens.*, vol. 9, no. 11, Oct. 2017, Art. no. 1110, doi: 10.3390/rs9111110.

[5] P. J. Zarco-Tejada, R. Diaz-Varela, V. Angileri, and P. Loudjani, "Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods," *Eur. J. Agronomy*, vol. 55, pp. 89–99, Apr. 2014, doi: 10.1016/j.eja.2014.01.004.

[6] P. J. Zarco-Tejada, A. Morales, L. Testi, and F. J. Villalobos, "Spatio-temporal patterns of chlorophyll fluorescence and physiological and structural indices acquired from hyperspectral imagery as compared with carbon fluxes measured with eddy covariance," *Remote Sens. Environ.*, vol. 133, pp. 102–115, Jun. 2013.

[7] D. Turner, A. Lucieer, M. McCabe, S. Parkes, and I. Clarke, "Pushbroom hyperspectral imaging from an unmanned aircraft system (UAS)—Geometric processing workflow and accuracy assessment," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W6, pp. 379–384, Aug. 2017.

[8] Z. Malenovský, A. Lucieer, D. H. King, J. D. Turnbull, and S. A. Robinson, "Unmanned aircraft system advances health mapping of fragile polar vegetation," *Methods Ecol. Evol.*, vol. 8, no. 12, pp. 1842–1857, Dec. 2017.

[9] T. Sankey, J. Donager, J. McVay, and J. B. Sankey, "UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA," *Remote Sens. Environ.*, vol. 195, pp. 30–43, Jun. 2017, doi: 10.1016/j.rse.2017.04.007.

[10] F. Tauro *et al.*, "Measurements and observations in the XXI century (MOXXI): Innovation and multi-disciplinarity to sense the hydrological cycle," *Hydrol. Sci. J.*, vol. 63, no. 2, pp. 169–196, Jan. 2018.

[11] J. Suomalainen *et al.*, "A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles," *Remote Sens.*, vol. 6, no. 11, pp. 11013–11030, Nov. 2014, doi: 10.3390/rs61111013.

[12] R. Hruska, J. Mitchell, M. Anderson, and N. F. Glenn, "Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle," *Remote Sens.*, vol. 4, no. 9, pp. 2736–2752, Sep. 2012.

[13] A. Lucieer, Z. Malenovský, T. Veness, and L. Wallace, "HyperUAS—Imaging spectroscopy from a multirotor unmanned aircraft system," *J. Field Robot.*, vol. 31, no. 4, pp. 571–590, Jul./Aug. 2014, doi: https://doi.org/10.1002/rob.21508.

[14] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, "Remote sensing image mosaicking: Achievements and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 8–22, Dec. 2019.

[15] R. Calderón, J. A. Navas-Cortés, C. Lucena, and P. J. Zarco-Tejada, "High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices," *Remote Sens. Environ.*, vol. 139, pp. 231–245, Dec. 2013.

[16] D. Schläpfer, M. E. Schaepman, and K. I. Itten, "PARGE: Parametric geocoding based on GCP-calibrated auxiliary data," *Proc. SPIE*, vol. 3438, pp. 334–344, 1998.

[17] J.-P. Ramirez-Paredes, D. J. Lary, and N. R. Gans, "Low-altitude terrestrial spectroscopy from a pushbroom sensor," *J. Field Robot.*, vol. 33, no. 6, pp. 837–852, Sep. 2016, doi: https://doi.org/10.1002/rob.21624.

[18] A. Habib, W. Xiong, F. He, H. L. Yang, and M. Crawford, "Improving orthorectification of UAV-Based push-broom scanner imagery using derived orthophotos from frame cameras," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 1, pp. 262–276, Jan. 2017.

[19] T. N. Shene, K. Sridharan, and N. Sudha, "Real-time SURF-based video stabilization system for an FPGA-driven mobile robot," *IEEE Trans. Ind. Electron.*, vol. 63, no. 8, pp. 5012–5021, Aug. 2016.

[20] S. Jeon, I. Yoon, J. Jang, S. Yang, J. Kim, and J. Paik, "Robust video stabilization using particle keypoint update and l_1-optimized camera path," *Sensors*, vol. 17, no. 2, Feb. 2017, Art. no. 337.

[21] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 2564–2571.

[22] Y. Angel, D. Turner, S. Parkes, Y. Malbeteau, A. Lucieer, and M. F. McCabe, "Automated georectification and mosaicking of UAV-based hyperspectral imagery from push-broom sensors," *Remote Sens.*, vol. 12, no. 1, Jan. 2020, Art. no. 34, doi: 10.3390/rs12010034.

[23] G. Hong and Y. Zhang, "Combination of feature-based and area-based image registration technique for high resolution remote sensing image," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 377–380.

[24] 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.*, 2019, pp. 9220–9223.

[25] R. Guerra *et al.*, "Optimal UAV movement control for farming area scanning using hyperspectral pushbroom sensors," in *Proc. XXXIVth Conf. Des. Circuits Integr. Syst.*, 2019, pp. 1–6, doi: 10.1109/DCIS201949030.2019.8959829.

[26] H. Yuan *et al.*, "Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models," *Remote Sens.*, vol. 9, no. 4, Mar. 2017, Art. no. 309.

[27] J. Wang, Y. Ge, G. B. M. Heuvelink, C. Zhou, and D. Brus, "Effect of the sampling design of ground control points on the geometric correction of remotely sensed imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 18, pp. 91–100, Aug. 2012, doi: http://dx.doi.org/10.1016/j.jag.2012.01.001.

[28] J. Hruška, L. Pádua, T. Adão, E. Peres, J. Martinho, and J. J. Sousa, "Target influence on ground control points (GCPs) identification in aerial images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 6487–6490.

[29] F. Agüera-Vega, F. Carvajal-Ramírez, and P. Martínez-Carricondo, "Assessment of photogrammetric mapping accuracy based on variation ground control points number using unmanned aerial vehicle," *Measurement*, vol. 98, pp. 221–227, Feb. 2017.

[30] G. Bradski, "The OpenCV library," *Dr. Dobb's J. Softw. Tools*, vol. 120, pp. 122–125, 2000, .

[31] R. Feng, Q. Du, X. Li, and H. Shen, "Robust registration for remote sensing images by combining and localizing feature-and area-based methods," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 15–26, May 2019.

[32] E. N. Mortensen, H. Deng, and L. Shapiro, "A SIFT descriptor with global context," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 1, pp. 184–190.

[33] E. Karami, S. Prasad, and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images," *CoRR*, 2017. [Online]. Avilable: https://dblp.org/rec/journals/corr/abs-1710-02726.bib

[34] M. Norouzi, D. J. Fleet, and R. R. Salakhutdinov, "Hamming distance metric learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1061–1069.

[35] M. Elbahnasawy *et al.*, "Multi-sensor integration onboard a UAV-based mobile mapping system for agricultural management," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2018, pp. 3412–3415.

[36] R. Ravi *et al.*, "UAV-based multi-sensor multi-platform integration for high throughput phenotyping," *Proc. SPIE*, vol. 11008, 2019, Art. no. 110080E, doi: 10.1117/12.2519190.

[37] T. Ivelja, B. Bechor, O. Hasan, S. Miko, D. Sivan, and A. Brook, "Improving vertical accuracy of UAV digital surface models by introducing terrestrial laser scans on a point-cloud level," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B1-2020, pp. 457–463, Aug. 2020, doi: 10.5194/isprs-archives-XLIII-B1-2020-457-2020.

[38] C. Van Veen, "UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA," *Remote Sens. Environ.*, vol. 195, pp. 30–43, Jun. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425717301578

[39] T. T. Sankey, J. McVay, T. L. Swetnam, M. P. McClaran, P. Heilman, and M. Nichols, "UAV hyperspectral and lidar data and their fusion for arid and semi-arid land vegetation monitoring," *Remote Sens. Ecol. Conservation*, vol. 4, no. 1, pp. 20–33, Mar. 2018.

[40] J. B. Sankey *et al.*, "Quantifying plant-soil-nutrient dynamics in rangelands: Fusion of UAV hyperspectral-LiDAR, UAV multispectral-photogrammetry, and ground-based lidar-digital photography in a shrub-encroached desert grassland," *Remote Sens. Environ.*, vol. 253, Feb. 2021, Art. no. 112223, doi: 10.1016/j.rse.2020.112223.

[41] J. Cao, W. Leng, K. Liu, L. Liu, Z. He, and Y. Zhu, "Object-Based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models," *Remote Sens.*, vol. 10, no. 1, Jan. 2018, Art. no. 89, doi: 10.3390/rs10010089.

[42] G. Bareth *et al.*, "Low-weight and UAV-based hyperspectral full-frame cameras for monitoring crops: Spectral comparison with portable spectroradiometer measurements," *Photogrammetrie-Fernerkundung-Geoinf.*vol. 2015, no. 1, pp. 69–79, Feb. 2015.

[43] G. Yang *et al.*, "The DOM generation and precise radiometric calibration of a UAV-mounted miniature snapshot hyperspectral imager," *Remote Sens.*, vol. 9, no. 7, Jul. 2017, Art. no. 642, doi: 10.3390/rs9070642.

[44] J. Yue *et al.*, "Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models," *Remote Sens.*, vol. 9, no. 7, Jul. 2017, Art. no. 708, doi: 10.3390/rs9070708.

[45] O. Nevalainen *et al.*, "Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging," *Remote Sens.*, vol. 9, no. 3, Mar. 2017, Art. no. 185, doi: 10.3390/rs9030185.

[46] X. Jiang, J. Ma, G. Xiao, Z. Shao, and X. Guo, "A review of multimodal image matching: Methods and applications," *Inf. Fusion*, vol. 73, pp. 22–71, Sep. 2021.

[47] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, no. 1, pp. 23–79, 2021.

[48] A. Habib, T. Zhou, A. Masjedi, Z. Zhang, J. E. Flatt, and M. Crawford, "Boresight calibration of GNSS/INS-Assisted push-broom hyperspectral scanners on UAV platforms," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1734–1749, May 2018.

[49] A. Lenz, H. Schilling, D. Perpeet, S. Wuttke, W. Gross, and W. Middelmann, "Automatic in-flight boresight calibration considering topography for hyperspectral pushbroom sensors," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 2981–2984.
[50] A. Habib, Y. Han, W. Xiong, F. He, Z. Zhang, and M. Crawford, "Automated ortho-rectification of UAV-Based hyperspectral data over an agricultural field using frame RGB imagery," *Remote Sens.*, vol. 8, no. 10, Oct. 2016, Art. no. 796, doi: 10.3390/rs8100796.
[51] M. Jaud *et al.*, "Direct georeferencing of a pushbroom, lightweight hyperspectral system for mini-UAV applications," *Remote Sens.*, vol. 10, no. 2, Feb. 2018, Art. no. 204, doi: 10.3390/rs10020204.
[52] J. P. Arroyo-Mora *et al.*, "Implementation of a UAV–Hyperspectral pushbroom imager for ecological monitoring," *Drones*, vol. 3, no. 1, Mar. 2019, Art. no. 12, doi: 10.3390/drones3010012.
[53] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
[54] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, "Locality preserving matching," *Int. J. Comput. Vision*, vol. 127, no. 5, pp. 512–531, May 2019, doi: 10.1007/s11263-018-1117-z.

Juan M. Jurado received the M.Sc. degree in computer engineering from the University of Jaén, Jaén, Spain, in 2017, the M.Sc. degree in high performance computing (HPC) from the University of A Coruña, A Coruña, Spain, in 2020, and the Ph.D. degree in computer science from the University of Jaén, Jaén, Spain, in 2020, writing his thesis on the "spectral characterizaiton and semantic segmentation of complex 3D models in natural environments".

He is currently a Professor with the Department of Computer Science, University of Jaén. His current research area focuses on the generation and processing of 3-D models, as well as the fusion of geometric, spatial, and spectral variables of real-world environments. The results of his research derive in significant advances in fields such as remote sensing, computer vision, and computer graphics through the development of novel methods for the integration of multisource data, the unsupervised classification of real scenarios, and the modeling of the material appearance.

Dr. Jurado was the recipient of the Award of the Best Master's Thesis of the University of Jaén in 2017.

Luís Pádua received the M.Sc. degree in computing engineering from the University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal, and the Ph.D. degree in informatics from the UTAD, with the thesis "automatic analysis of UAS-based multitemporal data as support to a precision agroforestry management system, in 2015 and 2021."

He is currently an External Research Collaborator with the Centre for Robotics in Industry and Intelligent Systems of Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal. He participated as a Research Assistant in projects related to the implementation of augmented reality platforms in cultural heritage sites and the application of unmanned aerial vehicles (UAVs) for vineyard monitoring. His thesis focused on the development of a decision support system for the agroforestry sector, based on the automatic analysis of UAV data obtained through different sensors in a multitemporal context. He has research interest in the areas of remote sensing applied to precision agriculture and forestry, unmanned aerial systems, data analysis, photogrammetric processing, digital image processing, computer vision, and machine learning.

Jonas Hruška received the bachelor's degree in geoinformatics from the Department of Geoinformatics, Palacký University Olomouc, Olomouc, Czechia, in 2014, and the master´s degree in geoinformatics from the Institute of Geoinformatics at VŠB, Technical University of Ostrava, Ostrava, Czechia, in 2016. He is currently working toward the Ph.D. degree in computer engineering with the University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal.

His bachelor thesis was focused on applying the stochastic Monte Carlo simulation in landscape development. In the master´s thesis, he was comparing data of remote sensing techniques for identification of slope deformation under vegetation cover with main concern to optical imagining using the UAV and laser scanning data. During the time of master´ thesis writing, he attended to a two-month long practical stay at UTAD where he gained the main knowledge about the UAS usage for science purposes. He is currently finishing his Ph.D. thesis focused on UAS and hyperspectral data exploring for agricultural purposes. He is also a Research Assistant with the UTAD.

Francisco R. Feito received the B.Sc. degree in mathematics from the Complutense University of Madrid, Madrid, Spain, and the Ph.D. degree in computer science from the University of Granada, Granada, Spain, in 1977 and 1995.

He was the Head of the Department of Computer Science, University of Jaén, Jaén, Spain, from 1993 to 1997 and 2003 to 2008, a Vice Chancellor in charge of studies and quality, from 1997 to 1999, and a Vice Chancellor in charge of research and international affairs with the University of Jaén, from 1999 to 2002. He is currently a Full Professor with the Department of Computer Science, University of Jaén. He has been teaching with the High Polytechnics Institute, Jaén, and the Head of the Department and the Graphics and Geomatics Research Group. His research interests include formal methods for computer graphics, geometric modeling, computational geometry, geographical information sciences, virtual archaeology, and precision agriculture.

Joaquim J. Sousa received the master's degree in surveying engineering from the University of Porto, Porto, Portugal, in 2001, working in the DEM extraction using radargrammetry, and the Ph.D. degree in surveying engineering from the University of Porto, Porto, Portugal (and TU Delft, Delft, The Netherlands), in 2010, writing his thesis on the "potential of integrating PSInSAR methodologies in the detection of surface deformation."

In 2010, he became an Assistant Professor and joined the Engineering Department, School of Science and Technology, Trás-os-Montes and Alto Douro, Vila Real, Portugal, as a Senior Lecturer in surveying engineering (including, among other topics: remote sensing, GIS, topography, technical drawing, and programming). He is the author of some national publications on GIS and topography. His current interest focuses on the use of unmanned aerial vehicles for forest and agriculture applications, especially with regard to image processing algorithms, using RGB, NIR, multispectral and hyperspectral, thermal and LiDAR sensors. In summary, precision agriculture and forestry, remote sensing, in-field sensing, and applied machine learning are among his main interests.