ISPRS Journal of Photogrammetry and Remote Sensing 134 (2017) 96–109

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Band registration of tuneable frame format hyperspectral UAV imagers in complex scenes

Eija Honkavaara [a,*], Tomi Rosnell [a], Raquel Oliveira [b], Antonio Tommaselli [b]

[a] Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute FGI, Geodeetinrinne 2, FI-02430 Masala, Finland
[b] Department of Cartography, São Paulo State University UNESP, Rua Roberto Simonsen 305, 19060-900 Presidente Prudente, Brazil

ARTICLE INFO

Article history:
Received 20 November 2016
Received in revised form 11 September 2017
Accepted 19 October 2017
Available online 6 November 2017

Keywords:
Hyperspectral imaging
UAV
Registration
Geometry
Photogrammetry

ABSTRACT

A recent revolution in miniaturised sensor technology has provided markets with novel hyperspectral imagers operating in the frame format principle. In the case of unmanned aerial vehicle (UAV) based remote sensing, the frame format technology is highly attractive in comparison to the commonly utilised pushbroom scanning technology, because it offers better stability and the possibility to capture stereoscopic data sets, bringing an opportunity for 3D hyperspectral object reconstruction. Tuneable filters are one of the approaches for capturing multi- or hyperspectral frame images. The individual bands are not aligned when operating a sensor based on tuneable filters from a mobile platform, such as UAV, because the full spectrum recording is carried out in the time-sequential principle. The objective of this investigation was to study the aspects of band registration of an imager based on tuneable filters and to develop a rigorous and efficient approach for band registration in complex 3D scenes, such as forests. The method first determines the orientations of selected reference bands and reconstructs the 3D scene using structure-from-motion and dense image matching technologies. The bands, without orientation, are then matched to the oriented bands accounting the 3D scene to provide exterior orientations, and afterwards, hyperspectral orthomosaics, or hyperspectral point clouds, are calculated. The uncertainty aspects of the novel approach were studied. An empirical assessment was carried out in a forested environment using hyperspectral images captured with a hyperspectral 2D frame format camera, based on a tuneable Fabry-Pérot interferometer (FPI) on board a multicopter and supported by a high spatial resolution consumer colour camera. A theoretical assessment showed that the method was capable of providing band registration accuracy better than 0.5-pixel size. The empirical assessment proved the performance and showed that, with the novel method, most parts of the band misalignments were less than the pixel size. Furthermore, it was shown that the performance of the band alignment was dependent on the spatial distance from the reference band.

1. Introduction

Hyperspectral imaging employs tens to hundreds of contiguous bands in order to accurately reconstruct the spectral signature of the target of interest (Goetz, 2009). Hyperspectral imagers with a variety of imaging principles have been operated from satellites and manned airborne platforms for decades (Schaepman, 2009). Whiskbroom imagers capture object spectra in pixels; examples of these sensors are the AVIRIS (Vane et al., 1993) and HyMap (Cocks et al., 1998). Pushbroom scanners, such as CASI (Babey and Anger, 1989) and the AISA series (Specim, 2017), capture spectra in lines. The use of sensors operating in the frame format principle (for example, those based on filter wheels or tuneable filters) have been rare due to the associated processing challenges (Schaepman, 2009).

Miniaturised hyperspectral sensor technology has developed rapidly in recent years, and the sensors are being implemented in small unmanned airborne vehicles (UAV). These novel technologies offer completely new opportunities for carrying out environmental remote sensing tasks. Several pushbroom hyperspectral sensors have been recently implemented in UAVs (Zarco-Tejada et al., 2012; Hruska et al., 2012; Büttner and Röser, 2014; Lucieer

* Corresponding author at: National Land Survey of Finland, Finnish Geospatial Research Institute FGI, Department of Remote Sensing and Photogrammetry, P.O. Box 84, FI-00521 Helsinki, Finland.
E-mail addresses: eija.honkavaara@nls.fi (E. Honkavaara), tomi.rosnell@nls.fi (T. Rosnell), r.alvoliveira@gmail.com (R. Oliveira), tomaseli@fct.unesp.br (A. Tommaselli).

https://doi.org/10.1016/j.isprsjprs.2017.10.014
0924-2716/

et al., 2014; Suomalainen et al, 2014). Researchers have also implemented point-based spectrometers in UAVs (Uto et al., 2013; Burkart et al., 2015). Recently, hyperspectral cameras operating in the frame format principle have entered the market (Mäkynen et al., 2011; Saari et al., 2011, 2013; Honkavaara et al., 2013; Mannila et al., 2014; Aasen et al., 2015).

When considering different sensing principles in UAV, the advantages of the frame imaging approach over the traditional pushbroom or whiskbroom scanning approaches include the possibility to collect image blocks with stereoscopic multiple object views and the geometric and radiometric constraints provided by the rigid rectangular image geometry (Honkavaara et al., 2013). Furthermore, georeferencing processing of 2D frame format sensors is simpler than that of pushbroom sensors that typically require expensive and heavy high-end direct georeferencing systems based on global navigation satellite system and inertial measurement unit (GNSS/IMU). Novel structure-from-motion (SFM) or bundle block adjustment technologies (Wu, 2013) and dense image matching technologies (Hirschmüller, 2005; Gruen, 2012) enable efficient and reliable image orientation and reconstruction of the object 3D structure based on frame-format images. In many applications, the 3D information is a significant feature parallel to the spectral information, for instance precision agriculture, vegetation monitoring, and forest measurements (Honkavaara et al., 2013; Aasen et al., 2015; Näsi et al., 2015). Commercially available frame-format hyperspectral imagers or sensors include the Rikola Hyperspectral Camera (Senop, 2017), the Cubert UHD 185-Firefly (Cubert, 2017) and the IMEC SM5X5 and SM4X4 (Imec, 2017).

When concerning mobile applications, the frame sensors can be further classified based on the imaging principle as those capturing registered spectral bands (snapshot imaging) or as those capturing non-registered bands (Aasen et al., 2015). The technologies collecting registered bands are based on mosaic filters, which reduce the spatial resolution, and the full resolution is provided by applying interpolation. Technologies that produce non-registered bands include techniques such as multiple cameras, filter wheels and tuneable filters. Different techniques could also be combined, for example, mosaic filters and tuneable filters. This study concerns band registration of the novel hyperspectral camera based on a tuneable Fabry-Pérot interferometer (FPI) (Mäkynen et al., 2011; Saari et al., 2011, 2013; Honkavaara et al., 2013). The FPI technology makes it possible to manufacture a lightweight, frame-format hyperspectral imager operating on the time-sequential principle, and it is also commercially available (Senop, 2017). The camera captures each individual hypercube by scanning the spectral range with different spectral settings within a short time interval, for example 1–2 s, by modifying the air gap between the FPI. As a result, the individual bands are 2D frame-format images, captured successively at each FPI gap size along the camera spectral range (Honkavaara et al., 2013; Oliveira et al., 2016). Thus, this technology will produce hypercubes with non-registered bands, when operated on a moving platform. More details of the technology and sensor used in this study are given in Sections 2.1 and 3.1. The above division is useful for mobile data capture; when operated in static mode, these techniques can also generate registered bands.

The image registration process is required in order to compensate for the band misalignment. Image registration is a technique to transform a target image into a reference image frame or a certain map projection (Dawn et al., 2010). Generally, image registration contains four main steps: (1) feature extraction, (2) feature matching, (3) image transformation, and (4) image interpolation or resampling (Dawn et al., 2010; Zitová and Flusser, 2003). Transformations can be two dimensional (such as rigid body, helmert, affine, polynomial or projective) or three dimensional, based on the collinearity model and accounting for the object 3D structure, i.e. the orthorectification process (Mikhail et al., 2001).

A previous study by Tommaselli et al. (2015) assessed empirically the potential of various 2D transformations in band registration of the time-sequential FPI camera images. The results showed that accuracy of registration decreased when the relative distance of the bands to the reference band increased or when the height variations in the object increased. Besides, the polynomial transformation outperformed the affine transformation. Results by Honkavaara et al. (2013) and Vakalopoulou and Karantzalos (2014) showed that the band registration of FPI images using feature-based matching and 2D image transformations provided good registration in flat agricultural scenes. Jhan et al. (2016) developed an approach utilising the relative calibration information and projective transformations for sensor systems with several rigidly integrated cameras, such as the MiniMCA lightweight camera, which is composed of six individual, integrated cameras. The approach was based on determining the relative orientations of individual cameras in the laboratory with respect to the master camera in the multi-camera system. The relative orientations of the master camera (red band) and an additional RGB camera were determined. The RGB camera was oriented with bundle-block adjustment, and the object-to-image transformations of the rest of the images were calculated based on relative orientations. An accuracy of 0.33 pixels was reported. Several researchers reported accuracies on the level of approximately 2 pixels when using 2D band registration approaches with the MiniMCA camera (Laliberte et al., 2011; Torres-Sanchez et al., 2013; Turner et al., 2014).

The objective of this investigation was to develop a method for accurate registration of non-registered hyperspectral data cubes captured using cameras based on tuneable filters. An important property of tuneable filter-based hyperspectral imaging is that the exterior orientation of each band in the image block is unique. Thus, each band presents its own central perspective projection of the object. If an object has 3D height variations, these will produce scale differences, which cannot be modelled accurately using 2D image transformations. This investigation presents and assesses an approach for determining accurate band registration accounting for the 3D object geometry that is determined by the structure-from-motion (Wu, 2013) and the dense digital image-matching technologies (Hirschmüller et al., 2005; Leberl et al., 2010; Rosnell and Honkavaara, 2012; Lisein et al., 2013; Eltner and Schneider, 2015).

Section 2 presents the FPI based hyperspectral imaging technology and the 3D approach for band registration. The method was tested using three forest image blocks. Section 3 presents the materials and methods, and Section 4 presents the results. Discussion is given in Section 5.

2. Theory of geometric processing of hyperspectral images based on a tunable filter

2.1. FPI camera technology

The hyperspectral camera technology developed by the VTT Technical Research Centre of Finland (VTT) (Mäkynen et al., 2011; Saari et al., 2011, 2013; Mannila et al., 2014) is based on a variable air gap FPI. When the FPI is placed in front of the sensor, the wavelength of the light passing the FPI is a function of the interferometer air gap. By changing the air gap, it is possible to acquire a new set of wavelengths. The final spectral response is dependent on the light passing the FPI and the spectral characteristics of the detector. The spectral bands can be selected according

to the requirements of the remote sensing task. In various implementations, three-colour (Saari et al., 2011, 2013; Honkavaara et al., 2013) or single-colour sensors have been used (Mannila et al., 2014; Oliveira et al., 2016; Senop, 2017). The spectral sensitivities have covered the visible- to near-infrared (VNIR) 400–1000 nm and short-wave infrared (SWIR; 1100–1600 nm) spectral ranges. The number of transmission peaks passing the FPI is 1 to 3; thus, exposure with a single gap width provides 1 to 3 different spectral bands when a three-colour sensor is used (Saari et al., 2011, 2013; Honkavaara et al., 2013). With a single-colour sensor, one spectral band is obtained for each air gap value.

In the UAV operation, the FPI camera is operated using photogrammetric principles, capturing image blocks with overlapping images according to the camera limits. During collection of each hyper cube, a predefined sequence of air gap values is applied to capture the full spectral range. Thus, the hyperspectral data cube is formed in the time-sequential imaging principle. During the data capture, the integration time, gain and FPI air gap information are recorded for each spectral band. FPI cameras are equipped with a GPS receiver that provides the exact time of the beginning of each data cube. Furthermore, the sensor electronics output the synchronisation pulse of each exposure.

2.2. Geometric characteristics of hyperspectral frame images captured using the time delay principle

When operating a hyperspectral frame imager based on a tuneable filter on a moving platform, each band in the data cube exposed to a different air gap value has a slightly different position and orientation. All these effects have to be taken into account in the post-processing step. In typical photogrammetric block imaging, the platform moves along an approximately linear track during the capture of a single hyperspectral cube, which results in a displacement of the bands along the same cube in the flight direction due to the forward movement. When assuming an ideal process, the forward movement will cause a shift between two bands a and b (ds_{ab}):

$$ds_{ab} = v_{ab} dt_{ab} \text{ [m]} \quad (1)$$

where v_{ab} is the flight speed, and dt_{ab} is the time difference between the bands. The shift can be scaled to the image coordinates by multiplying by the scale factor $sf = c/FH$, where c is the principal distance and FH is the flying height. The shift is variable because the flight speed is changing all the time according to wind conditions and other factors. Furthermore, the platform can be non-stable, causing relative orientation differences between the bands along all axes, which can result in mismatches in any direction; platform stabilisation will minimise these impacts. When operating the sensor in block imaging mode, these effects are expected to be minor in comparison to the mismatches of bands in the flight direction due to the linear movement along the flight trajectory. This investigation focuses data capture from an airborne platform flying on a linear trajectory; thus, the main part of the displacement is due to the forward motion along the trajectory.

In images captured with the perspective projection principle, the object height variations from the elevation above or below a particular projection plane causes a so-called relief displacement effect, i.e., a shift in an object's image position (Fig. 1a; Mikhail et al., 2001). For vertical or near-vertical photography, the shift occurs radially from or towards the nadir point (Fig. 1b). In the time-sequential spectral images, the relief displacement of a certain object will be different at different bands (a, b) due to the position difference caused by the time difference (dt_{ab}) between the bands. Relief displacements in the bands a and b can be calculated as follows for an object of height dZ (Fig.1a):

$$dX_a = dZ\ X_a/(FH - dZ) = dZ \tan \alpha \text{ [m]} \quad (2)$$

$$dX_b = dZ\ X_b/(FH - dZ) = dZ(X_a + v_{ab} dt_{ab})/(FH - dZ) \text{ [m]}$$

where X_a and X_b are the distances of the object from the vertical projection of the perspective centre on the ground in two bands (i.e. distance from the nadir point); α is the view angle from the vertical direction to the point; dZ is the height of the object from the reference plane and FH is the flight height that can be considered to remain constant during the two bands of the cube. Thus, the difference in relief displacements is

$$dX_{ab} = dX_b - dX_a = dZ\ v_{ab} dt_{ab}/(FH - dZ). \text{ [m]} \quad (3)$$

The objective in sensor manufacturing is to minimise the time difference between adjacent bands, for example, in current implementations, the values are 0.075 s or 0.033 s; the time difference between the first and last bands depends on the number of airgaps applied, being approximately 1.8 s or 1.0 s, respectively (Honkavaara et al., 2013; Oliveira et al., 2016). These cube times have produced good-quality hypercubes in our UAV flight missions, which have usually been carried out using flight speeds of 3 m/s to 5 m/s, with wind speeds of less than 6 m/s, and using flight altitudes of 100–140 m. For example, when the time difference dt_{ab} is 1 s between the first and the last band in the cube, the flight height is 100 m, and the flight speed is 5 m/s, the difference in relief displacements is about 1.25 m for an object with a height of 20 m (e.g. a treetop), and 0.05 m for an object with a height of 1 m (Fig. 1c). For the ground sample distance (GSD) of 10 cm, this will cause discrepancies of 12.5 and 0.5 pixels, respectively. Displacement differences can be decreased by increasing the flight height and decreasing the flight speed, the time difference between the bands, and the height differences in the object (Eq. (3)). Currently, due to legislation, the maximum allowable flight altitudes are typically 100–150 m; thus, the technology has some limitations with respect to allowable flight speeds and, in these cases, it is applicable for slow platforms, such as multicopters.

If the flight speed is too high, different bands do not see the same areas of the object. Even when the overlap is acceptable, for example, the first and last captured bands are overlapping e.g. 90% or more, the occlusions, due to perspective imaging, will appear. The visibility of a specific point in the image is dependent on the view angle (α) and the object steepness, and can be analysed using visibility analysis techniques (Mikhail et al., 2001; Habib et al., 2007). dX in Fig. 1 also gives the occlusion of the specific object; the dimension of the occlusion can be calculated using Eq. (2). The maximum occlusions thus appear at large view angles and high objects. The specific challenge in the case of FPI imaging is the possibility that the same object details are not visible in different bands due to occlusions. The flight parameters have to be selected so that an adequate representation of the object is obtained.

In order to align the bands perfectly, the registration has to be performed using accurate object height information and a rigorous model describing the imaging event. If registration is performed using 2D transformations, the differences in relief displacements in different bands will appear as image coordinate errors and will cause internal discrepancies in the hyperspectral data cubes. This displacement is serious particularly in scenes with large height differences, such as forests. The following section develops a rigorous approach for band registration based on a collinearity model, which overcomes the problems due to variations of terrain displacements.

2.3. 3D band registration method

In rigorous processing, the object scale differences have to be accounted for by utilising an object 3D model. Geometric transfor-

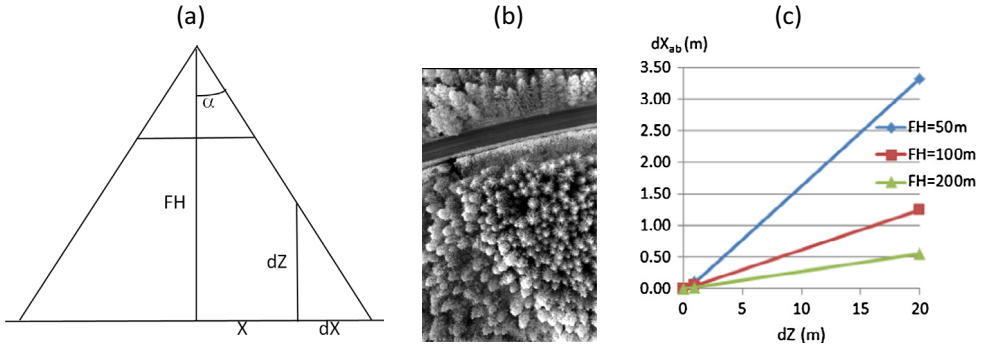

Fig. 1. (a) Illustration of relief displacement (dX) for an object at a height of dZ located at distance X from the nadir position. (b) The relief displacement of trees appears as a radial shift away from the image centre. (c) Impact of platform movement during hypercube acquisition in relief displacement differences in two bands (dX_{ab}) for an object with a height of dZ; 1 s time difference, flight speed 5 m/s.

mation is thus performed using a collinearity model augmented with image distortion parameters (Fraser, 1997):

$$x - x_0 - \delta x = -c\frac{m_{11}(X-X_0)+m_{12}(Y-Y_0)+m_{13}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)}$$
$$y - y_0 - \delta y = -c\frac{m_{21}(X-X_0)+m_{22}(Y-Y_0)+m_{23}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \quad (4)$$

where x and y are the observed image point coordinates; X, Y, and Z are the ground coordinates of the same point; m_{ij} are the elements of the rotation matrix; X_0, Y_0, and Z_0 are the ground coordinates of the camera's perspective centre; x_0 and y_0 are the principal point coordinates; c is the camera's principal distance; and δx and δy are the effects of lens and image distortions causing deviation from the collinearity model. The image perspective centre coordinates and the image rotations compose the image exterior orientation parameters (EOPs), and the principal distance, principal point coordinates and the image distortion parameters form the interior orientation parameters (IOPs).

Registration based on the collinearity equations requires determination of the EOPs of each image, the IOPs of the cameras and the 3D geometry of the object. The fundamental methods for determining EOPs of airborne images are the bundle-block adjustment and direct georeferencing. When using bundle-block adjustment, the IOPs and EOPs are adjusted to provide the best fit of images in the image block and with the object reference. Direct georeferencing is based on determining the image positions and rotations during the flight by means of GNSS and IMU, and then performing the final georeferencing using this information, object 3D information and camera calibration information. Direct georeferencing-based methods are efficient, but generally the remaining small uncertainties in the direct orientation measurement, inaccuracies in synchronisation and timing information and instability of camera calibration, datum errors, etc., will cause an unknown amount of georeferencing error (Cramer and Stallmann, 2001; Cramer, 2002), which can be disturbing in applications that require accuracies on the level of pixel size. The 3D geometry of the object can be determined accurately using dense image matching (Hirschmüller, 2005; Gruen, 2012).

In this study, a rigorous method based on bundle-block adjustment is developed. When concerning hyperspectral data cubes with non-registered bands, EOPs of each band in each data cube can be determined using the bundle-block adjustment. However, this can lead to a time-consuming calculation process, for example, if more than 100 bands need to be included in the block adjustment calculation. One potential approach is to calculate accurate EOPs for a limited number of bands and interpolate the EOPs of the rest of the bands from the trajectory. This investigation adapts and improves the previous method by calculating the orientations of image blocks of a selected number of reference bands using the bundle-block adjustment and then using a rigorous space resection method for the remaining unoriented bands. The resection means determination of an image's EOPs with respect to an object-space coordinate system (Mikhail et al., 2001; Luhmann et al., 2014). The resection is a special case of bundle-block adjustment, where the object is known and the image orientation parameters need to be solved. A minimum of three non-collinear control points (CPs) are required for a resection solution of a frame image, assuming that the IOPs are known. In the case of band registration, the process thus fits the unoriented bands optimally to the reference bands and object.

Principles of the proposed registration method are described below; the implementation used in this study is described in detail in Section 3.2:

1. Selection of reference bands. The selection of reference bands should take into account the temporal (positional) and spectral characteristics of the data cubes, since both the displacement of the bands and the spectral response have an impact in the band-matching process. For example, reference bands could be taken in the beginning, middle and end of the data cube and, furthermore, from each major spectral range (e.g., blue, green, red and near-infrared). An appropriate number of reference bands should be considered for each application; typically 1–3 reference bands have been used.
2. Determination of IOPs and EOPs of one or more reference bands and the object model. For the metric cameras, the IOPs are specific to a camera. They can be determined in a laboratory calibration process, or they can be calculated using a self-calibrating bundle-block adjustment. In the case of an FPI camera, the previous study by Oliveira et al. (2016) showed that the IOPs determined for one of the bands fitted well for the other bands that were captured on the same image plane using the optical path, where the FPI air gap was the only component that changed. Separate IOPs were required for the bands on two separate image planes of an FPI camera system based on two single-colour CMOS sensors and a beam-splitting component. In the proposed approach, the bundle-block adjustment is used to calculate the IOPs and EOPs of the reference bands, as well as the tie points' 3D ground coordinates (sparse point cloud). The image blocks are formed from the reference band images of all hypercubes. After a self-calibrating bundle-block adjustment, dense point clouds are calculated by image matching (Hirschmüller, 2005; Gruen, 2012). The next steps run separately for each band in each hypercube.

3. Feature extraction. The 3D coordinates estimated for the tie points in the bundle-block adjustment, i.e. the sparse point cloud, are used as control points (CP) for the resection solution. They are the features used in image registration. CPs are back-projected to the reference band and to the unoriented band images using the collinearity model. For the reference band, the solved IOPs and EOPs provide accurate mapping between the sparse point cloud and image coordinates. For the unoriented bands, the interpolated orientations from the flight trajectory are used, which provides the approximate image coordinates of the tie points.
4. Feature matching. The accurate image coordinates of the reference features are searched in the unoriented bands using the image matching process. Area-based matching, based on the correlation coefficient function (Kraus, 2007), is used in this investigation to find the best match among the tie points in the reference bands and the unoriented bands. The performance of the image matching is tuned by the size of the search window, the size of the template (feature) window and the limit for an acceptable correlation coefficient. Selection of suitable parameters can be based on empirical study or automatic and adaptive selection methods (Kanade and Okutomi, 1994; Hosni et al., 2013; Yang et al., 2014).
5. Image transformation. The collinearity model (Eq. (4)) is used as the image transformation. The EOPs of unoriented images are calculated by the space resection technique (Mikhail et al., 2001; Luhmann et al., 2014). The principles of IOPs are presented in Step 2. The sparse point cloud calculated in Step 2 provides the CP object coordinates and the feature matching (Step 4) provides the CP image coordinates. In the processing, all points in the sparse point cloud be used in the calculation, typically several thousands of points in each image. Our approach selects part of the points to the space resection in order to reduce the computational cost of the calculation. To obtain good point distribution, we divide the image area into a grid, for example with 10×10 grid cells, and select a certain number of points in each grid cell, typically 1–10 points. These parameter values (grid cells and number of points per cell) can be changed if needed. The impact of the number of points is one of the research questions in this investigation. Outliers are always present in the automatic image matching procedures, thus means to limit number of outliers and identify the outliers are implemented in the processing.
6. Output product resampling. The resampling process calculates the hyperspectral output products that have aligned bands. In this study, either orthophoto mosaics (Mikhail et al., 2001) or point clouds with spectral data can be provided. The spectral values for each XYZ-point or orthophoto mosaic pixel are obtained from the images using the collinearity Eq. (4) by using the 3D object model (DSM or point cloud) and the EOPs and IOPs of the images. Typically, the most nadir method is used, which means that the spectral value is taken from the image where the view angle to the object is the smallest. In the output product calculation, the radiometric processing is also of importance as discussed by Honkavaara et al. (2012 and 2013).

2.4. Uncertainty propagation in 3D band matching

The uncertainties in the parameters in Eq. (4) cause a geometric uncertainty in the reflectance output products and band alignment. Uncertainty estimation in photogrammetry is based on the variance–covariance propagation techniques (Mikhail, 1976; Luhmann et al., 2014). In the following, we first consider the general aspects of the uncertainty estimation in photogrammetry, and then we will consider the principles of the uncertainty estimation in the hypercube band registration process.

The functional model in photogrammetry is based on the collinear geometrical relationship between the object coordinate and the image point and is extended by some additional parameters (Eq. (4)). The linearized functional model is:

$$\hat{l} = l + v = A\hat{x}, \tag{5}$$

where $\hat{l}$ is the vector of adjusted observations and l is the vector of observations, A (design matrix) provides the functional relationships between the parameters in the model and contains the derivatives with respect to model parameters, v is the vector of the residuals of the observations, and $\hat{x}$ is the vector of unknowns. (Mikhail, 1976; Luhmann et al., 2014.)

The least squares technique provides estimates of the unknowns:

$$\hat{x} = (A^T W_{ll} A)^{-1} A^T W_{ll} l \tag{6}$$

where W_{ll} is the weight matrix of the observations. Residuals are calculated as:

$$v = A\hat{x} - l \tag{7}$$

$\hat{\sigma}_0$ is the a posterior standard deviation:

$$\hat{\sigma}_0 = \sqrt{v^T W_{ll} v/(n-u)} \tag{8}$$

where n is the number of observations and u is the number of unknowns. The variance covariance matrix of unknowns is:

$$\hat{\Sigma}_{xx} = \hat{\sigma}_0^2 (A^T W_{ll} A)^{-1}. \tag{9}$$

When calculating derived products, such as hyperspectral mosaics or hyperspectral point clouds, the uncertainties of the parameters from the previous phases are further propagated to the output products. The derived coordinate (z) is obtained as:

$$z = g(\hat{x}). \tag{10}$$

The uncertainties of derived coordinates can be determined using the variance–covariance propagation technique (Mikhail, 1976):

$$\hat{\sigma}_{zz}^2 = J_{zx} \hat{\Sigma}_{xx} J_{zx}^T, \tag{11}$$

where J_{zx} is the Jacobian matrix of $g(\hat{x})$ and $\hat{\Sigma}_{xx}$ is the variance–covariance matrix of the parameters of the model.

The uncertainty propagation is implemented in the georeferencing process of time-sequential hypercubes (Section 2.3) as follows. The exterior orientations of reference bands are determined using a bundle-block adjustment. In the second phase, the orientations of the unoriented bands are determined by resection, relatively to the reference bands. In both cases, the uncertainties of the parameters can be calculated using Eq. (9). In the image resampling process, the coordinates of each orthophoto mosaic pixel or point cloud point are calculated in the oriented images using the collinearity equations (Eq. (4)), and the grey values are interpolated from surrounding pixels. Uncertainty of the position of a calculated point can be determined using the variance propagation (Eq. (11)). In the equations, normally distributed random errors with a zero mean are assumed. However, especially the DSM can contain outliers due to matching errors. The uncertainty assessment, based on standard deviations, does not account for these errors. The outliers in the object 3D model will cause different errors in bands collected at different times; the differences will follow principles in Eq. (3).

3. Material and methods

3.1. Data sets

The procedure was tested using three forest image blocks. Two of the data sets were captured in Finland in the Vesijako research forest area of the Natural Resources Center Finland, in the municipality of Padasjoki in southern Finland (approximately 61°24′ N and 25° 02′ E). The area, named v02, was dominated by birch (*Betula pendula*), whereas area v06 was dominated by spruce (*Picea abies*). The Vesijako datasets were used previously in a tree species classification study by Nevalainen et al. (2017); for that study altogether eight image blocks were successfully processed with the proposed method. The third area was located in the Wytham Woods 51°46′33″N, −1°20′12″E, Oxford, UK. Wytham Woods are an ancient semi-natural mixed woodland owned by the University of Oxford (Wytham Woods, 2017). The dominating tree species is maple (*Quercus robur*). The Finnish areas were typical managed forest areas with separable, relatively narrow tree tops. Tree densities were on average 898 and 1381 trees per ha for the test plots v02 and v06, respectively. The Wytham area had continuous tree cover with mature broad-crowned trees.

A hexacopter UAV with Tarot 960 foldable frame was used as the platform (Fig. 2). The autopilot was a Pixhawk equipped with Arducopter 3.15 firmware. The FPI camera prototype 2012b belonging to the Finnish Geospatial Research Institute (FGI) was used to capture FPI hyperspectral images in the range of 500–900 nm (Honkavaara et al., 2013; Saari et al., 2013). It is equipped with custom optics, with a principal distance of 10.9 mm and an aperture of 2.8. The camera has a CMOSIS CMV4000 RGB image sensor, with a global electronic shutter; single exposure provides 1–3 simultaneous bands. There is 0.075 s between exposures with consecutive air gap values. For the prototype 2012b 1–24 airgap values can be selected; the time lapse from the first to the last exposure in an air gap sequence with 24 gap values (single hypercube) is 1.8 s. The image size is 1024 × 648 pixels, with a pixel size of 11 μm. The field of view (FOV) is ±18° in the flight direction, ±27° in the cross-flight direction, and ±31° at the format corner. The entire camera system weighs less than 700 g. An RGB camera, the Samsung NX 300, was used to support the geometric processing. The camera had a 16 mm lens and an image size of 23.5 mm by 15.7 mm; the FOV was ±26° in the flight direction, ±36° in the cross-flight direction, and ±45° at the format corner. The cameras were rigidly installed on the landing gear of the UAV.

The image blocks consisted of six (Vesijako v02), four (Vesijako v06) and ten (Wytham w01) flight lines (Fig. 3, Tables 1 and 2). The flight speed varied between 3.8 m/s and 4.6 m/s. The nominal flight height in each area was 90 m. For the approximate tree height of 20 m, the distance to object was about 70 m. The resulting Ground Sample Distance (GSD) was about 9 cm at ground level and about 7 cm at treetop for the FPI camera. For the RGB camera, the GSD was about 2.5 cm at ground level and about 2 cm at the tree tops. The data sets in Finland were captured under cloudy conditions, whereas in Wytham the weather was cloudless and sunny. Image overlaps were considered separately at treetop and at ground level (Table 2). They were the lowest for the test area v02. At treetop level, the forward and side overlaps of FPI images were 53% and 43%, respectively. For the RGB data set, the poorest overlaps were 69% and 61%, respectively. The overlaps in the RGB data sets were acceptable in all cases, whereas for the FPI data sets, the overlaps were lower than typically recommended for the structure-from-motion-based algorithms. The RGB images were used to support the processing because of the low overlaps for FPI images.

Fig. 2. The hexacopter UAV, equipped with the FPI camera and the Samsung NX 300 consumer RGB camera.

Table 3 and Fig. 4 show the spectral and temporal settings for the image data sets. Temporal settings included the time difference (dt) and the positional difference (ds) in metres and in pixels from the start of the hyper cube. The positional difference from the first band was, at the maximum time lapse of 1.8 s, approximately 7.2 m in the object space, and 100 and 80 pixels in the image space for tree tops and ground objects, respectively (Fig. 4). The maximum difference in the object's image positions due to scale differences between different bands was thus approximately 20 pixels (Eq. (3)). The differences were approximately 5% smaller for the Wytham dataset because of the lower flight speed. The airgap sequence provided single band for bands 14–25 and 2 simultaneous bands for bands 1–13 and 26–38. Fig. 5 shows examples of three non-registered bands of one data cube.

3.2. Geometric processing

The processing method described in Sections 2.3 and 2.4 were implemented in the photogrammetric environment at the FGI. The processing of the reference bands (including the determination of EOPs and IOPs and the calculation of sparse and dense point clouds) was carried out using the commercial photogrammetric software Agisoft PhotoScan Professional (AgiSoft LLC, St. Petersburg, Russia). The rest of the processing steps (including the band matching, orthophoto mosaic generation, uncertainty estimation and error estimation) have been implemented as C++ software at the FGI.

PhotoScan performs photo-based 3D reconstruction based on feature detection and dense matching. Its excellent performance has been validated in previous studies (Eltner and Schneider, 2015; Näsi et al., 2015). We used a simultaneous geometric processing for the RGB images and three bands in Vesjako from the FPI cubes (see Table 4). In Wytham data, we used the RGB images and single band from the FPI images. The advantages of using the RBG and FPI images in the same orientation process included higher overlaps between the images in the block in comparison to the case with only one band. This is important if overlaps of individual images are low, as was the case with the data sets in Finland. There was also improved accuracy because the RGB images had better spatial resolution, which improved the accuracy of the feature-matching based orientation. The numbers of FPI hyper cubes per block varied from 39 to 304 (Table 2). In the PhotoScan processing, the quality was set to "high", which indicated that full resolution images were used in the processing. The setting for the number of tie points per image was 1000 and a lens calibration was performed simultaneously using the self-calibration. The initial processing provided image orientations and sparse point clouds in the internal coordinate system of the software. An automatic outlier removal was performed on the basis of the residuals (10% of the points with the largest residuals were removed), as well as

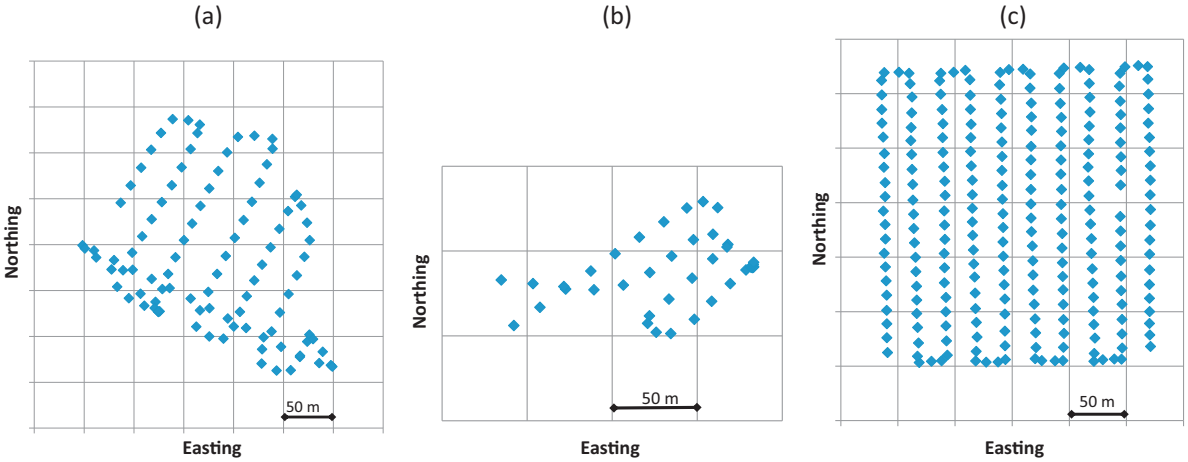

Fig. 3. Image blocks (a) Vesijako (v02), (b) Vesijako 06 (v06), and (c) Wytham woods (w01).

Table 1
Information of the image data sets for each study area.

Area	Date	Time GPS)	Speed (m/s)	Weather	Object	Sun Elev.; Az	Median irradiance	Exposure (ms)
v02	26.6.14	12:09 to 12:22	4.6	Cloudy	Birch	47.4; 219.6	4427	12
v06	25.6.14	12:14 to 12:24	4.3	Cloudy	Spruce	47.0; 221.3	3773	10
w01	30.6.15	09:46 to 10:04	3.8	Sunny	Deciduous	51.6; 125.2	10817	5

Table 2
Information about image blocks are calculated with respect to the ground level and the tree top level. N images: number of images; FH: flight height; fw: forward overlap; sl: side overlap.

Block	N images		Block properties at ground level					Block properties at treetop level				
	FPI	RGB	FH (m)	FPI GSD (cm)	fw;sl (%)	RGB GSD (m)	fw;sl (%)	FH (m)	FPI GSD (m)	fw;sl (%)	RGB GSD (m)	fw;sl (%)
v02	98	176	88	8.8	65;57	2.4	77;70	67	6.7	53;43	1.8	69;61
v06	39	76	94	9.4	78;79	2.5	85;86	73	7.3	71;73	2.0	80;81
w01	304	257	90	9.0	77; 72	2.4	92;80	70	7.0	71;64	1.9	90;75

Table 3
Spectral and temporal properties of hypercubes. L0: peak wavelength, FWHM: full width at half maximum; time difference (dt), movement of the platform in metres (ds) from the start of the hyper cube. Bands 1–13 have simultaneous bands in NIR range (26–38).

L0 (nm): 507.60, 509.50, 514.50, 520.80, 529.00, 537.40, 545.80, 554.40, 562.70, 574.20, 583.60, 590.40, 598.80, 605.70, 617.50, 630.70, 644.20, 657.20, 670.10, 677.80, 691.10, 698.40, 705.30, 711.10, 717.90, 731.30, 738.50, 751.50, 763.70, 778.50, 794.00, 806.30, 819.70, 833.70, 845.80, 859.10, 872.80, 885.60
FWHM (nm): 11.2, 13.6, 19.4, 21.8, 22.6, 20.7, 22.0, 22.2, 22.1, 21.6, 18.0, 19.8, 22.7, 27.8, 29.3, 29.9, 26.9, 30.3, 28.5, 27.8, 30.7, 28.3, 25.4, 26.6, 27.5, 28.2, 27.4, 27.5, 30.5, 29.5, 25.9, 27.3, 29.9, 28.0, 28.9, 32.0, 30.8, 27.9
dt (s): 0.9, 0.975, 1.05, 1.125, 1.2, 1.275, 1.35, 1.425, 1.5, 1.575, 1.65, 1.725, 1.8, 0, 0.075, 0.15, 0.225, 0.3, 0.375, 0.45, 0.525, 0.6, 0.675, 0.75, 0.825, 0.9, 0.975, 1.05, 1.125, 1.2, 1.275, 1.35, 1.425, 1.5, 1.575, 1.65, 1.725, 1.8
Vesijako v02 and v06: ds (m): 3.6, 3.9, 4.2, 4.5, 4.8, 5.1, 5.4, 5.7, 6, 6.3, 6.6, 6.9, 7.2, 0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3, 3.3, 3.6, 3.9, 4.2, 4.5, 4.8, 5.1, 5.4, 5.7, 6, 6.3, 6.6, 6.9, 7.2
Wytham w01: ds (m): 3.4, 3.7, 4.0, 4.3, 4.6, 4.8, 5.1, 5.4, 5.7, 6.0, 6.3, 6.6, 6.8, 0.0, 0.3, 0.6, 0.9, 1.1, 1.4, 1.7, 2.0, 2.3, 2.6, 2.9, 3.1, 3.4, 3.7, 4.0, 4.3, 4.6, 4.8, 5.1, 5.4, 5.7, 6.0, 6.3, 6.6, 6.8

reconstruction uncertainty (10% of the points with the largest uncertainty were removed). In addition, some outliers were manually removed from the sparse point cloud (points on the "sky" and below the ground). Ground Control Points (GCPs) with an accuracy of 0.5–1 m were used in order to transform the image orientations to the object coordinate system. The outputs of the process were the image EOPs and the IOPs of each camera. Dense point clouds were generated only using the RGB images in Vesijako and using the RGB and FPI images in Wytham. This process was carried out using two-times downsampled images (setting "high" in Photoscan). The filtering options "moderate" and "mild" were used to eliminate outliers in Vesijako and Wytham, respectively.

PhotoScan processing provided orientation for the reference bands (Table 4). The EOPs of the other bands were recovered using the proposed method (Section 2.3); single reference band was used in the final space resection calculation (band 4 in Vesijako and band 6 in Wytham). The points used in band matching were based on the sparse point cloud (tie points) provided by the PhotoScan, which provided approximately 1000 points per image. In order to control the number of the CPs per images, we first divided the image area considering a grid with 6×10 cells. The impact of the number of points were tested using for each grid cell 1, 5 or 10 CPs, which provided approximately 60, 300 or 600 points, respectively, for the resection process of each image.

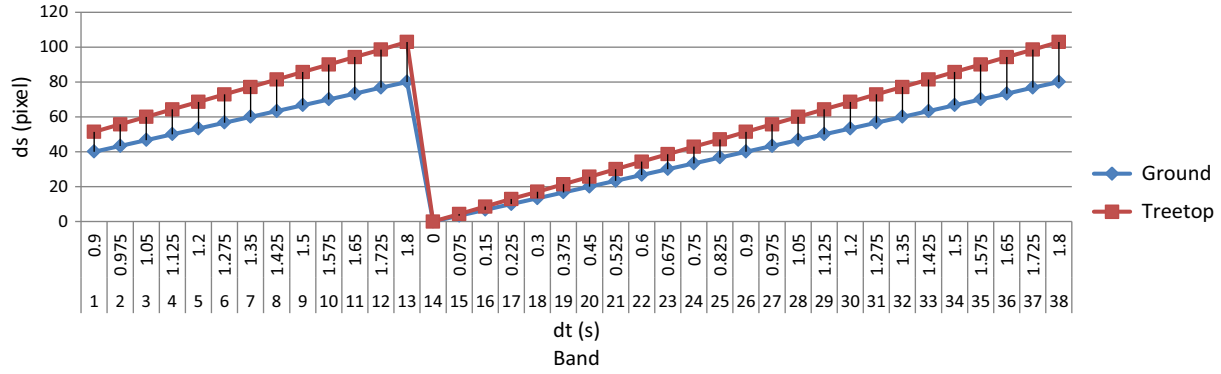

Fig. 4. Visualisation of the difference in relief displacement of different bands (dX_{ab}) with respect to the first band during the exposure at ground and at treetop level for the Vesijako dataset. Time difference is from the beginning of the hypercube.

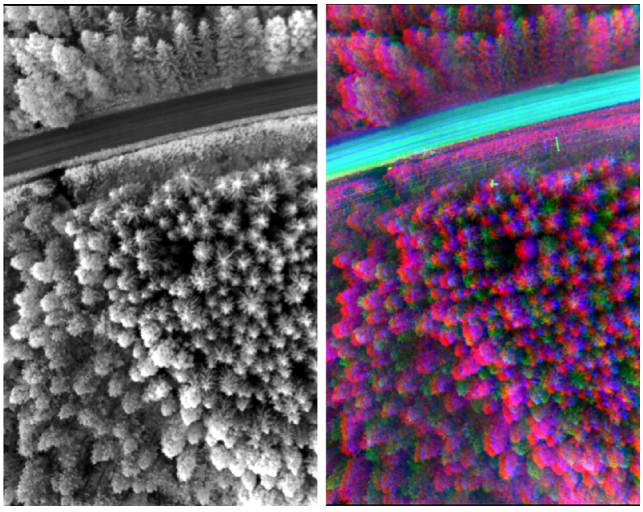

Fig. 5. Individual band 33 (L0 = 819.7 nm) in the NIR region and the non-registered band composition with bands 4 (L0 = 520.8 nm), 14 (L0 = 605.7 nm) and 33 (L0 = 819.7 nm) from the Vesijako v06 area. Distances to the first exposure of the hypercube were in time 1.125 s, 0 s and 1.425 s and at tree top level (GSD 0.07 cm) 64.3, 0 and 81.4 pixels, respectively.

Table 4
Reference bands used in block adjustment. The bands in parentheses are simultaneous bands to the bands that were adjusted in the block adjsutment. L0: peak wavelength, dt: time difference to the cube start.

Block	Band id	L0 (nm)	dt (s) (to cube start)
v02, v06	4 (29), 12 (37), 16	520.8, 590.4 (763.7), 630.7 (872.8)	1.125, 1.725, 0.150
w01	6 (31)	537.4 (794.0)	1.275

During the image matching, the size of the search window was 200 × 200 pixels (18 m × 18 m on ground level) and the size of the template (feature) window was 15 × 15 pixels (1.3 m × 1.3 m on ground level). The correlation coefficient threshold was set to 0.7. In order to avoid outliers, a forward/backward matching strategy was used. This means that the matched points in an unoriented image are matched back to the oriented image. The limit for the difference in position in backward matching images was three pixels in row and column directions. During the space resection process, a recursive outlier removal scheme was applied to remove possible matching errors. The resection statistics were evaluated to further identify unsuccessful calculations: a minimum of 10 CPs (for setting with 60 CPs per image) or 30 CPs (for setting with 300 or 600 CPs per image) were required for each image; the limit of 30 μm was set for acceptable $\hat{\sigma}_0$. These settings were selected empirically, and our studies with numerous data sets from forests with similar GSDs have shown that the same parameters have been suitable for all the datasets.

Finally, orthophoto mosaics of each band were generated with a 10 cm GSD. Also, hyperspectral point clouds were calculated with 10 cm point interval. The EOPs, IOPs and DSM from the previous processing were used as the *a priori* data in the output production calculation. Grey values were interpolated using the nearest neighbour interpolation from the images where the view angle to the object was the smallest, i.e., using the most nadir method.

3.3. Accuracy assessment

A theoretical performance assessment was carried out by analysing the variance propagation (Section 2.4). In order to obtain the estimated uncertainty of the band alignment, we introduced the IOPs and EOPs of the reference bands and the CPs as errorless parameters and calculated the full variance-covariance matrix for the EOPs calculated by the space resection. Then, we used the full variance-covariance matrix and error propagation technique (Eq. (11)) to calculate the uncertainties of the geometric transformations used in the mosaic and point cloud calculation.

An empirical performance assessment of the band alignment in the ortophoto mosaics was carried out using image matching of each band to the reference band. In this process, a template window, 15 × 15 pixels in size (1.3 m × 1.3 m), was extracted from the image under evaluation, and its position was searched in the reference (search) image in an image window, 9 m × 9 m in size. Discrepancies were calculated and classified as 0, 1, 2, 3, etc., pixels in X and Y directions. Proportions (in per cent) belonging to each class were calculated.

The IOPs and EOPs and the dense point clouds were calculated by the PhotoScan, which did not provide *a posteriori* variance–covariance matrix of the estimated parameters. Thus, we could not use uncertainties of exterior orientations to assess the total uncertainty. The final estimate of the spatial errors of orthophoto mosaics could be calculated by assuming the uncertainties due to block adjustment and due to band alignment independent:

$$\hat{\sigma}_x = \sqrt{\hat{\sigma}^2_{x\_ref} + \hat{\sigma}^2_{x\_bp}}, \tag{12}$$

where $\hat{\sigma}_{x\_ref}$ is uncertainty of the reference band coordinate transformation and $\hat{\sigma}_{x\_bp}$ is the uncertainty due to ground point back projection (Eq. (11)).

The national airborne scanning (ALS) data by the National Land Survey of Finland (NLS) was used as the geometric reference for evaluating the geometric quality of phtogrammetric processing of data sets captured in Finland (NLS, 2016). The minimum point density of the NLS's ALS data is half a point per square metre, and the elevation accuracy of the points in well-defined surfaces is 15 cm. The horizontal accuracy of the data is 60 cm. The point cloud has an automatic ground classification. The ALS data used in this study was collected on 12 May 2012. The ALS data has much lower resolution than the photogrammetric DSM, but can be used to obtain a general view of the success and deformations of the photogrammetric block.

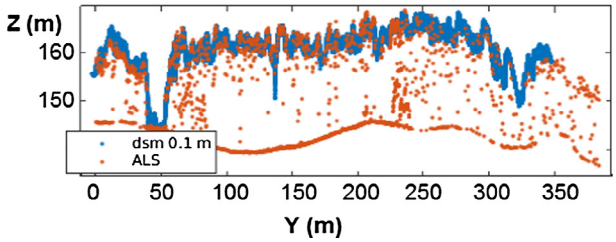

Fig. 6. Example of comparison of airborne laser scanning (ALS) and photogrammetric point cloud (dsm 0.1 m) in the Vesijako v02 test area.

4. Results

4.1. Processing of reference bands

Block processing indicated good orientation accuracy with the root mean square errors (RMSE) of residuals (reprojection errors), varying between 0.5 and 1.1 pixels (Table 5). For Vesijako v02 and v06 blocks, the results of RGB block processing are also given because the DSMs were calculated using this data. For the data from Finland, the ALS-based reference model could be used as the reference in an approximate DSM quality assessment (Fig. 6). Visually, the dense point clouds had good quality when comparing to the ALS reference model. The UAV photogrammetric surfaces followed the structure that was visible in the ALS point cloud, individual trees were well aligned, and minor differences in canopy top level and ground level were observable. As anticipated, the photogrammetric surface did not reach the ground level as often as ALS. For the Wytham data, the visual assessment indicated that the image orientations and the DSM were of good quality. The quality of the georeferencing of reference bands could be considered as sufficient for processing at an accuracy level of a single pixel (10 cm).

4.2. Assessment of theoretical uncertainty of band registration

EOPs of the bands not inserted in the bundle adjustment were calculated using the proposed resection method (Sections 2.3 and 3.2). Calculation was successful and only a few images failed. These problems appeared during the take-offs and landings of the UAV and in turns in the block, as the simple correlation method used in matching could not take care of the great rotation differences in images. The RMSEs of image residuals after back projection were 5–15 μm (0.5–1.4 pixels). The further the matched band was from the reference band, the larger the RMSEs were (Fig. 7).

The orthophoto mosaics were calculated with a 10 cm GSD. The theoretical uncertainty estimates of point back projection of orthophoto mosaics showed a clear dependency on the distance from the reference band (Fig. 8) and consistently to the back projection residuals (Fig. 7). Resections were calculated with nominal numbers of 60, 300 and 600 CPs per image. Theoretical estimates indicated that as the number of points used in the resection increased, the standard deviations of the back-projected points decreased, i.e. the orientation accuracy improved. The largest standard deviations appeared for the smallest numbers of points and the largest distances from the reference band, which was consistent with expectations. The largest standard deviations were about 0.5 pixels for the area v06 and 0.3 pixels or less for other data sets. The simultaneous bands in the NIR range had slightly poorer accuracy than the bands in the red and green spectral range; this was due to slightly reduced accuracy of matching of NIR bands to the reference band in the green spectral range.

4.3. Empirical assessment of output product quality

Examples of non-registered hypercubes and corresponding registered mosaics are shown in Fig. 9a and b, the full mosaics of each test area are shown in Fig. 9c-e, and an example of a hyperspectral point cloud from Wytham with spectral information is given in Fig. 9f. Visual assessment showed that the quality was good in central areas of the blocks, where image overlaps were good and view angles were small. Some distortions and relief displacements appeared in the border areas where only oblique view data was available, and a few pixel mismatches appeared in areas where the DSM did not accurately follow the canopy surface (Eq. (3)). Maximum view angles to the object can be calculated based on the image overlaps and image FOV. The computational maximum view angle is approximately (100-overlap)/2% of the FOV. In the data sets, in the flight direction, the maximum view angles were 2-3°, and in the cross-flight direction, the maximum view angles were 3-6°. This provides maximum occlusions for the approximate tree height of 20 m of 0.7–1.1 m in the flight direction and of 1.0–2.0 m in the cross-flight direction. In practice, the edges of trees are not sharp, thus the actual occlusions were smaller. Furthermore, the canopy was so dense that the ground level was visible only in openings. In the areas where the ground was visible, the registration quality was good also at the ground level.

Discrepancies of registration of each ortophoto mosaic band was assessed empirically by matching the bands to the reference band (Fig. 10). Discrepancies were classified to 1 pixel (Fig. 10a) and 0 pixels (Fig. 10b), indicating either a discrepancy of 1 pixel (or less) or an exact match. The percentages of appearance of each class were calculated. The impact of distance from the reference band is clearly visible in the plots. The results were quite similar in x and y directions. In all the cases, the majority (93–99%) of the differences were classified as 1 pixel or less. The portions of errorless matches were 60–99% and 40–95% for the Vesijako and

Table 5
Block statistics are given for the cases with only FPI images and for cases with combined FPI and RGB data sets.

Block	N GCPs	N images	Tie points	Projections	Reprojection error (pix)	Points/m^2
v02 fpi+rgb	4	469	82450	281795	0.64	
v02 rgb	4	176	35705	94471	0.73	140
v06 fpi+rgb	3	193	18 088	79149	0.91	
v06 rgb	3	76	10291	39750	1.10	130
w01 fpi+rgb	6	881	123039	329595	0.54	629

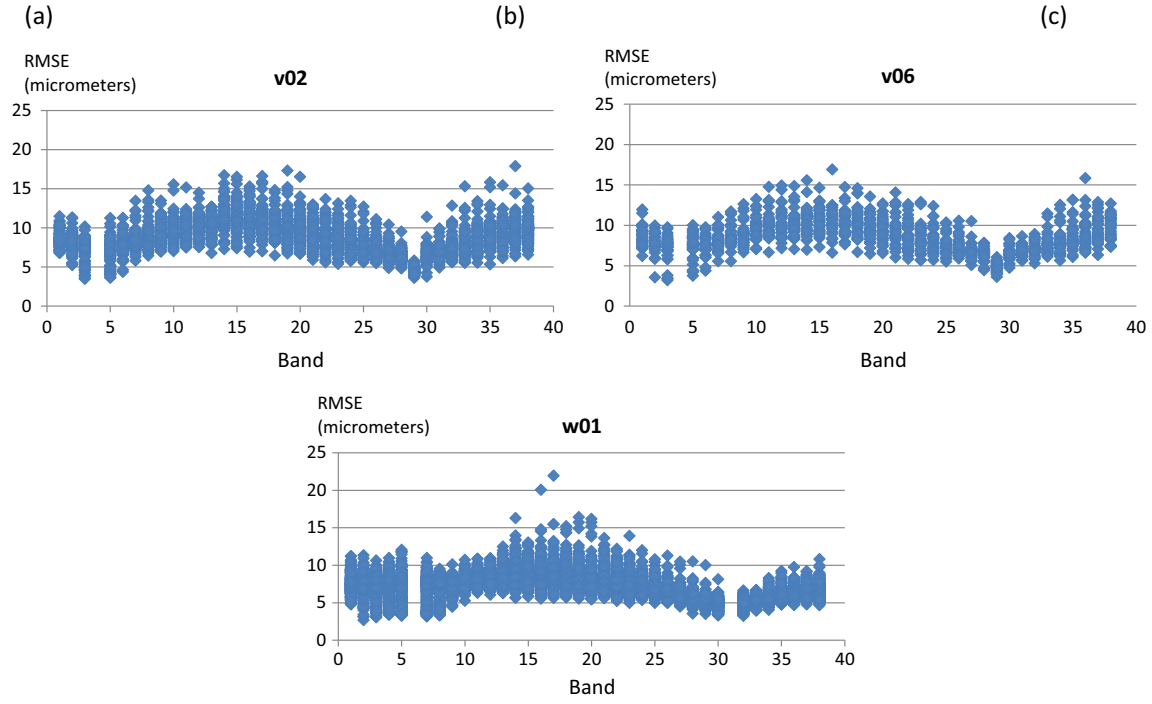

Fig. 7. RMSEs of residuals after the resection for blocks (a) v02, (b) v06 and (c) w01 for each band and hypercube.

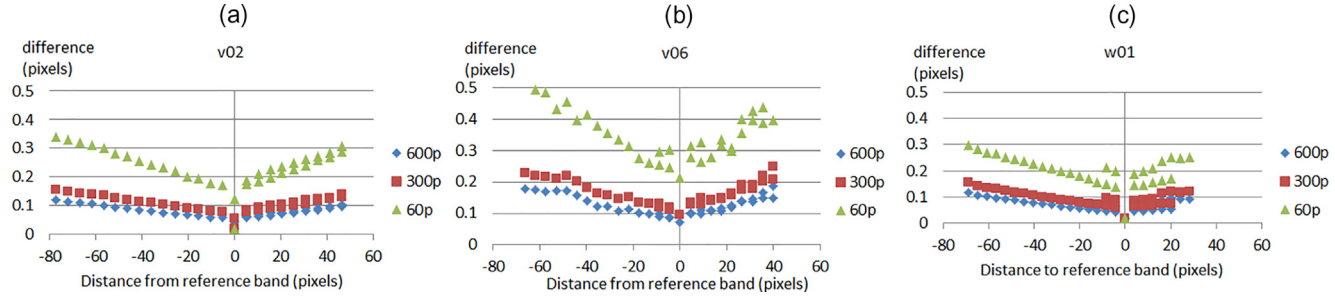

Fig. 8. (a) Theoretical backprojection uncertainties (RMSE of calculated uncertainties of each back projected mosaic coordinate) for test areas (a) Vesijako v02, (b) Vesijako v06 and (c) Wytham w01 for each band presented as the function of distance from the reference band. Different colours indicate cases with 60, 300 and 600 control points per image (60p, 300p, 600p).

Wytham data sets, respectively. Discrepancies were calculated for different cases with nominal numbers of 60, 300 and 600 CPs per image. However, they were similar in each case; thus, only the case with 600 points per image is shown. Examples of distributions of the discrepancies are shown for each data set in Fig. 10(c) for band 9 that has approximately a 0.3–0.4 s time difference to the reference band.

5. Discussion

Hyperspectral imaging based on tuneable filters is highly interesting technology for UAV-based remote sensing. The misalignment of individual bands is a serious challenge in utilisation of this technology. Our results showed, theoretically and empirically, that the proposed approach provided accurate band registration for the time-sequential hyperspectral images in a forested scene. The method is based on the accurate orientation of a few reference bands using a rigorous bundle-block adjustment, subsequent registration of the un-oriented bands using resection that utilises the reference band processing results (orientations, camera calibration, and sparse point clouds), and final generation of output products using dense photogrammetric point cloud. It was shown that registration accuracy between ortophoto mosaics of individual bands at a one-pixel level or better was possible when good georeferencing and high-quality DSM were available. In order to obtain good results, the entire measurement setup has to be designed properly, in particular, the ground control points or direct georeferencing solutions and block overlaps.

Previous research results have already proven the feasibility of the FPI hyperspectral imaging technology in agricultural applications (Honkavaara et al., 2013; Vakalopoulou and Karantzalos, 2014) and forestry applications (Näsi et al., 2015). In those studies, the band registration was based on 2D image transformations. This investigation provided the theoretical framework to understand the geometric characteristics of the tuneable frame format imaging and developed an efficient, rigorous 3D approach to perform the band registration. The accurate band registration is a prerequisite for detailed analysis, such as, forest analysis at individual tree level, precision agriculture or object characterisation in urban areas. For object scenes having great height differences with respect to the flight height, the 3D-based approaches are required in order to obtain registration accuracy better than pixel size as was shown in this study. For flat targets, 2D-matching-based approaches also

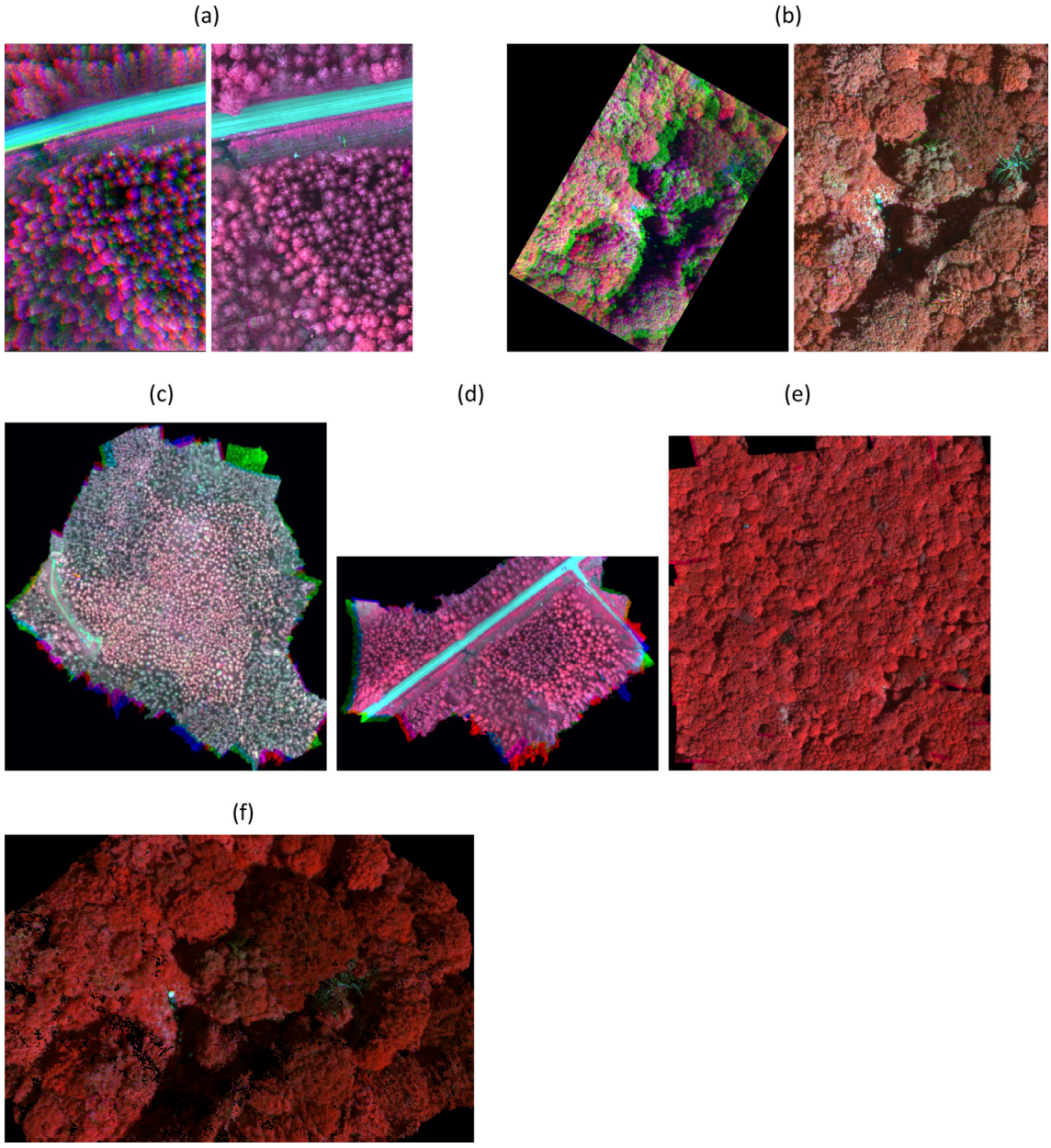

Fig. 9. (a) Non-registered and registered datacube from Vesijako v06 area (bands 4, 14 and 33). (b) Non-registered and registered datacube from Wytham w01 area (bands 6, 14 and 33). Examples of ortophoto mosaics (c) Vesijako v02, (d) v06 and (e) Wytham w01. (f) A hyperspectral point cloud from Wytham w01.

provide accurate registration. Accurate band registration is required, because the geometric inaccuracies in the band registration can lead to challenges in interpretation phase and to incorrect association of spectral information to object, which may distort and complicate the analysis procedures.

Resection is typically not the preferred solution in photogrammetric processing lines, especially due to the high cost of capturing control points and due to the risk of potential inconsistencies between individual photos. The presented approach resolves these deficiencies by integrating the resection process tightly to the bundle-block adjustment. The bundle-block adjustment provides a large number of control points for the resection process with minor additional cost; and when using large numbers of control points, excellent consistency is obtained between individual images and bands. Implementing the calculation is simple, because the approximate orientations are quite accurately known on the basis of the trajectory data of the oriented bands. The advantage of using a few selected reference bands in the bundle-block adjustment instead of using all bands is that the huge growth of the adjustment problem is avoided. The individual band resections also allow a simple way to parallelise the processing.

RGB and FPI camera images were processed simultaneously in the bundle-block adjustment. The RGB images provided high spatial resolution, a wide field of view and good overlaps, improving the georeferencing quality. When using the RGB image block, it was possible to use lower image overlaps for the FPI images that had a smaller field of view and would have required more flight lines in order to allow high-accuracy photogrammetric reconstruction. The view angles to the object were less than 10° in most nadir image mosaics due to a narrow field of view in the FPI images, which provided an acceptable level of occlusions in the image mosaics. Jhan et al. (2016) also used a higher resolution RGB sensor

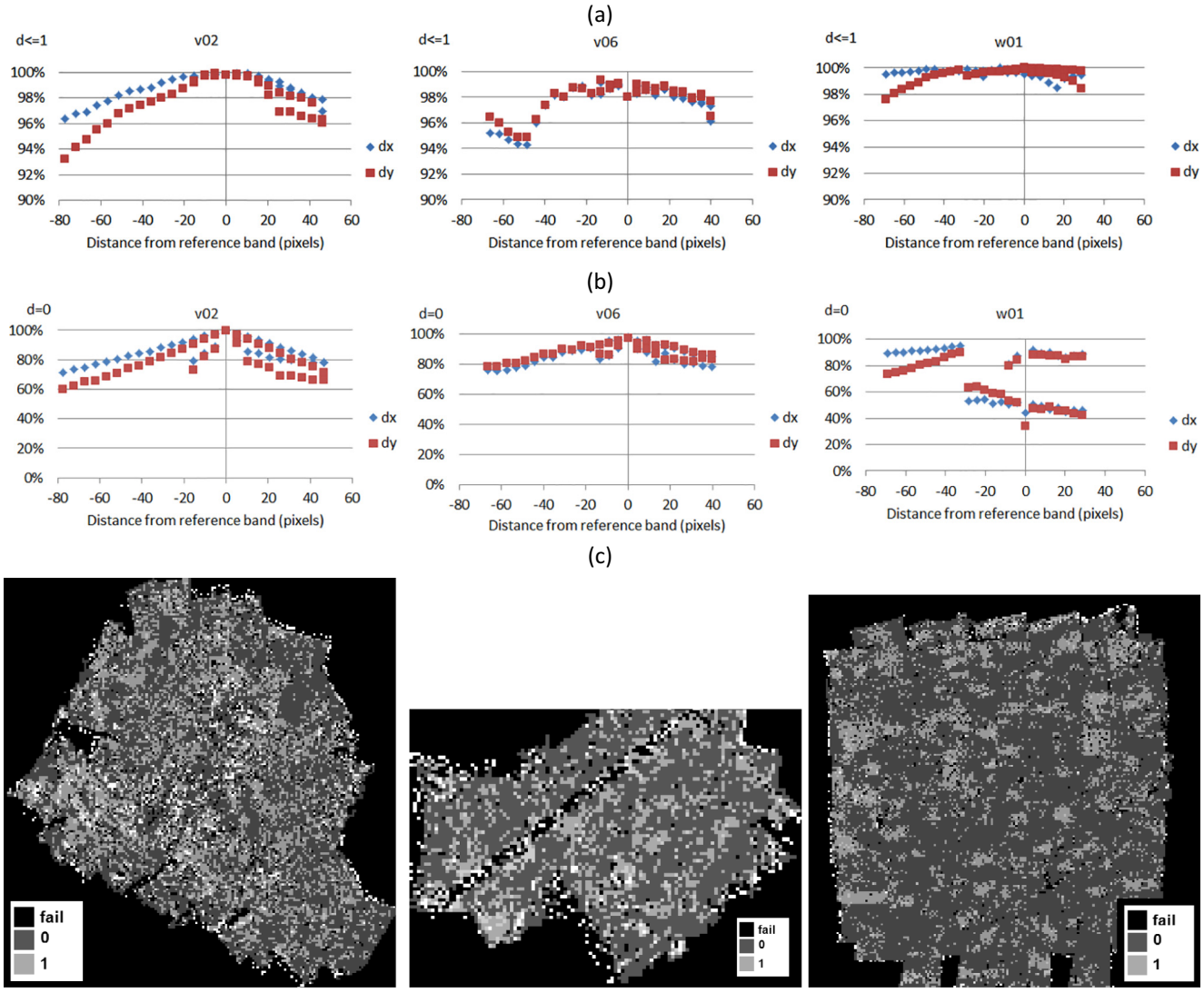

Fig. 10. Empirical error assessment by correlation matching, test areas from left, Vesijako v02 and v06, and Wytham w01. (a) The percentages of discrepancies <=1. (b) The percentages of exact matches (discrepancy = 0). (c) Discrepancy in each matched point between band 9 and reference band (time difference approximately 0.3–0.4 s).

to support the georeferencing with the MiniMCA multispectral camera. It is worth noting that, with appropriate block setup, the entire processing could be performed by using the FPI images only.

Image matching was used to localise the tie points in unoriented bands. The success of this process is dependent on the image texture and spectral characteristics of the objects. In a forested scene, the single reference band in the green spectral range provided accurate results, since relative spectral characteristics of the object in the spectral range of 500–900 nm were quite similar across each band inside a matching template area. The texture in the template area was composed by the shape of the tree crowns and shades in the crowns. In contrast, for instance, in agricultural areas, both ground and vegetation impact the texture of the template windows. In this situation, the bare ground appears brighter than the vegetation in the red and green bands; whereas, in the NIR range, the performance is the opposite. Thus, it can be expected that the matching, based on a grey value correlation, will not be successful in that case. The ideal band combinations and reference bands need to be considered separately for different applications. If necessary, more than one reference band is used. Furthermore, in texture-less scenes, such as large water areas, the approach can be unsuitable. For these cases, approaches based on direct georeferencing should be developed. In further developments, improving the speed of the feature-matching method will be desirable.

The FPI camera prototype 2012b was used in this study. It differs from the commercially available FPI cameras by having only a single optical path and detector; the commercial FPI cameras have two optical paths due to the beam splitter (Tommaselli et al., 2016; Oliveira et al., 2016; Senop, 2017). For the multiple path system, it is necessary to use a separate interior orientation for each individual sensor whereby each band can be matched to the selected reference band. A further difference from commercial cameras is the time differences of individual bands (0.075 s), which were larger than the time differences in the commercial versions of the FPI camera (0.033 s). Due to the faster image acquisition, the flight speed can be increased to provide similar spatial distances within data cubes, as in this study.

In order to obtain good object geometric reconstruction accuracy, it is crucial to use image blocks with large overlaps. The recommended overlaps are dependent on the object; for forested scenes, it is typically recommended to use greater than 70% side overlaps and 80% forward overlaps. Also, the sensor's field of view

should be accounted for when selecting overlaps; in this investigation, viewing angles in most-nadir orthophoto mosaics were less than 10°. Suitable view angles are important to avoid occlusions and to minimise impacts of bidirectional reflectance effects (Honkavaara et al., 2013). Good image quality (high signal-to-noise-ratio) is essential in order to obtain good 3D geometric object reconstruction by image matching.

Test areas in this investigation were fairly homogeneous forested scenes, with birches, conifers and semi-natural mixed, deciduous forest captured with similar flight parameters. Novel hyperspectral imaging technologies are of interest also in other types of complex scenes, such as urban scenes, tropical forest, mountains or glaciers. Previous studies have shown that the geometric processing based on the structure-from-motion-technique or bundle block adjustment, and dense DSM generation have excellent performance in a variety of objects if the scene includes enough features, for example, in tropical forested areas (Berveglieri et al., 2016), rugged terrain (Jaud et al., 2016) and even for the more homogeneous glaciers areas (Dall'Asta et al., 2017; Gindraux et al., 2017). Urban areas also provide suitable features for SFM and the dense image matching process (Nex and Remondino, 2014; Haala and Cavegn, 2016); the effect of the occlusions due to the buildings can be minimised for instance using low flight speed and high overlaps. Thus, the proposed registration approach is expected to be functional also for various complex environments.

In the future, we will further improve the processing of tuneable filter hyperspectral imaging. It is necessary to study and adopt the processing methodologies in different environments, such as agricultural scenes and the previously mentioned complex scenes. Furthermore, approaches for texture-less scenes, such as water areas or plain snow surface, should be developed. In these cases, direct georeferencing is an attractive approach, and its use in the context of tuneable filter frame format cameras should be investigated further. We expect that the further development of georeferencing methods will improve the overall georeferencing accuracy, which will also improve the registration quality of the tuneable filter images.

6. Conclusions

The present study developed a solution for one of the greatest bottlenecks in hyperspectral frame format imaging based on tuneable filters, namely a rigorous and efficient approach for band registration in complex 3D scenes in the case of UAV imaging. The approach utilises bundle-block adjustment/structure-from-motio n-based image orientation and subsequent spatial resection in calculating accurate exterior orientation parameters for each band. Further steps include generation of hyperspectral ortophoto mosaics and/or point clouds utilising the accurate object geometric reconstruction based on dense image matching. The method is rigorous and uncertainty propagation can be integrated into the method. The method is based on image matching, which is feasible for the processing of hypercubes that have suitable textures. A theoretical analysis showed that the discrepancies between the bands increased when the spatial difference to the reference band increased. In the evaluated cases, the discrepancies (RMSE) were always less than 0.5 pixels, and typically less than 0.3 pixels, assuming that the object height model was accurate. Errors in the height model will cause larger discrepancies in the output products. Empirical analysis proved that the accuracy of the band registration was 1 pixel or better. In the test data sets, 93% or more of the registration errors between the reference band and the matched band were 1 pixel or less.

Acknowledgements

The research carried out in this study was financially supported by the Academy of Finland (Project No. 273806) and the Finnish Funding Agency for Innovation, Tekes – Finland (Project No. 2208/31/2013). We express our sincere gratitude to Teemu Hakala for the UAV image data capture and to Niko Viljanen for the image orientation and digital surface model generation.

References

Aasen, H., Burkart, A., Bolten, A., Bareth, G., 2015. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. ISPRS J. Photogram. Rem. Sens. 108, 245–259. https://doi.org/10.1016/j.isprsjprs.2015.08.002.

Babey, S.K., Anger, C.D., 1989. A compact airborne spectrographic imager (CASI). In: Proc. IGARSS '89 and 12th Canadian Symposium on Remote Sensing, Vancouver, Canada, 10–14 July, 2, pp. 1028–1031. doi: 10.1109/IGARSS.1989.579067.

Berveglieri, A., Tommaselli, A.M.G., Imai, N.N., Ribeiro, E.A.W., Guimaraes, R.B., Honkavaara, E., 2016. Identification of successional stages and cover changes of tropical forest based on digital surface model analysis. IEEE J. Select. Top. Appl. Earth Observ. Rem. Sens. 9 (12), 5385–5397. http//dx.doi.org/10.1109/JSTARS.2016.2606320.

Burkart, A., Aasen, H., Alonso, L., Menz, G., Bareth, G., Rascher, U., 2015. Angular dependency of hyperspectral measurements over wheat characterized by a novel UAV based goniometer. Remote Sens. 7, 725–746. https://doi.org/10.3390/rs70100725.

Büttner, A., Röser, H.-P., 2014. Hyperspectral remote sensing with the UAS "Stuttgarter Adler" – system setup, calibration and first results. Photogram. Fernerkundung Geoinform. 2014, 265–274. https://doi.org/10.1127/1432-8364/2014/0217.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., Shields, T., 1998. The HyMap airborne hyperspectral sensor: the system, calibration and performance. In: Proc. 1st EARSeL Workshop on Imaging Spectroscopy, Zurich, Paris, pp. 37–43 (6–8 October).

Cramer, M., Stallmann, D., 2001. On the use of GPS/inertial exterior orientation parameters in airborne photogrammetry. In: Proc. ISPRS Workshop-High Resolution Mapping from Space, Hannover, Germany, pp. 32–44. 19–21 September.

Cramer, M., 2002. Experiences on operational gps/inertial system calibration in aiborne photogrammetry. J. GIS-Geoinformationssysteme 14 (6), 37–42.

Cubert, 2017. Cubert GmbH http://cubert-gmbh.de, (accessed 26 August, 2017).

Dawn, S., Saxena, V., Sharma, B., 2010. Remote sensing image registration techniques: a survey. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D., Meunier, J. (Eds.), Image and Signal Processing. Springer, Berlin, Heidelberg, pp. 103–112.

Dall'Asta, E., Forlani, G., Roncella, R., Santise, M., Diotri, F., di Cella, U.M., 2017. Unmanned aerial systems and DSM matching for rock glacier monitoring. ISPRS J. Photogram. Rem. Sens. 127, 102–114. https://doi.org/10.1016/j.isprsjprs.2016.10.003.

Eltner, A., Schneider, D., 2015. Analysis of different methods for 3D reconstruction of natural surfaces from parallel-axes UAV images. Photogram. Rec. 30 (151), 279–299. https://doi.org/10.1111/phor.12115.

Fraser, C.S., 1997. Digital camera self-calibration. ISPRS J. Photogram. Rem. Sens. 52 (4), 149–159. https://doi.org/10.1016/S0924-2716(97)00005-1.

Gindraux, S., Boesch, R., Farinotti, D., 2017. Accuracy assessment of digital surface models from unmanned aerial vehicles' imagery on glaciers. Remote Sens. 9 (2), 186. https://doi.org/10.3390/rs9020186.

Goetz, A., 2009. Three decades of hyperspectral remote sensing of the Earth: a personal view. Remote Sens. Environ. 113, S5–S16. https://doi.org/10.1016/j.rse.2007.12.014.

Gruen, A., 2012. Development and status of image matching in photogrammetry. Photogram. Rec. 27 (137), 36–57. https://doi.org/10.1111/j.1477-9730.2011.00671.x.

Haala, N., Cavegn, S., 2016. High density aerial image matching: state-of-the-art and future prospects. Int. Arch. Photogram. Rem. Sens. Spatial Inform. Sci. 41 (B4), 625–630. https://doi.org/10.5194/isprs-archives-XLI-B4-625-2016.

Habib, A.F., Kim, E.M., Kim, C.J., 2007. New methodologies for true orthophoto generation. Photogram. Eng. Rem. Sens. 73 (1), 25–36 https://dx.doi.org/10.14358/PERS.73.1.25.

Hirschmüller, H., 2005. Accurate and eficient stereo processing by semi-global matching and mutual information. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2, pp. 807-814. https://doi.org/10.1109/CVPR.2005.56.

Honkavara, E., Hakala, T., Saari, H., Markelin, L., Mäkynen, J., Rosnell, T., 2012. A process for radiometric correction of UAV image blocks. Photogram. Fernerkundung Geoinform. 2012 (2), 115–127.

Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., Litkey, P., Mäkynen, J., Pesonen, L., 2013. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. Remote Sens. 5 (10), 5006–5039. https://doi.org/10.3390/rs5105006.

Hosni, A., Bleyer, M., Gelautz, M., 2013. Secrets of adaptive support weight techniques for local stereo matching. Computer Vis. Image Underst. 117 (6), 620–632. https://doi.org/10.1016/j.cviu.2013.01.007.

Hruska, R., Mitchell, J., Anderson, M., Glenn, N., 2012. Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. Remote Sens. 4 (12), 2736–2752. https://doi.org/10.3390/rs4092736.

Imec, 2017. https://www.imec-int.com (accessed 26 August, 2017).

Jaud, M., Passot, S., Le Bivic, R., Delacourt, C., Grandjean, P., Le Dantec, N., 2016. Assessing the accuracy of high resolution digital surface models computed by PhotoScan® and MicMac® in sub-optimal survey conditions. Remote Sens. 8 (6), 465. https://doi.org/10.3390/rs8060465.

Jhan, J.-P., Rau, J.-Y., Huang, C.-Y., 2016. Band-to-band registration and ortho-rectification of multilens/multispectral imagery: A case study of MiniMCA-12 acquired by a fixed-wing UAS. ISPRS J. Photogram. Remote Sens. 114, 66–77. https://doi.org/10.1016/j.isprsjprs.2016.01.008.

Kanade, T., Okutomi, M., 1994. A stereo matching algorithm with an adaptive window: Theory and experiment. IEEE Trans. Pattern Anal. Mach. Intell. 16 (9), 920–932. https://doi.org/10.1109/34.310690.

Kraus, K., 2007. Photogrammetry, Geometry from Images and Laser Scans. Walter de Gruyter, Berlin, New York.

Laliberte, A.S., Goforth, M.A., Steele, C.M., Rango, A., 2011. Multispectral remote sensing from unmanned aircraft: image processing workflows and applications for rangeland environments. Remote Sens. 3 (11), 2529–2551 https://dx.doi.org/doi:10.3390/rs3112529.

Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S., Wiechert, A., 2010. Point clouds: Lidar versus 3D vision. Photogram. Eng. Rem. Sens. 76 (10), 1123–1134 https://dx.doi.org/10.14358/PERS.76.10.1123.

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., Lejeune, P., 2013. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. Forests 4 (4), 922–944 https://doi.org/doi:10.3390/f4040922.

Lucieer, A., Malenovský, Z., Veness, T., Wallace, L., 2014. HyperUAS-imaging spectroscopy from a multirotor unmanned aircraft system. J. Field Robot. 31 (4), 571–590 https://doi.org/doi:10.1002/rob.21508.

Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2014. Close Range Photogrammetry and 3D Imaging. Walter de Gruyter GmbH & Co. KG, Göttningen, Germany.

Mannila, R., Holmlund, C., Ojanen, H., Näsilä, A., Saari, H., 2014. Short-wave infrared (SWIR) spectral imager based on Fabry-Perot interferometer for remote sensing. In: SPIE Proc. 9241, 92411M. https://doi.org/10.1117/12.2067206.

Mikhail, E.M., 1976. Observations and Least Squares. University Press of America Inc., New York.

Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to Modern Photogrammetry. John Wiley & Sons Inc., New York.

Mäkynen, J., Holmlund, C., Saari, H., Ojala, K., Antila, T., 2011. Unmanned aerial vehicle (UAV) operated megapixel spectral camera. In: SPIE Proc. 8186, 81860Y. 10.1117/12.897712.

Nevalainen, O., Honkavaara, E., Tuominen, S., Viljanen, N., Hakala, T., Yu, X., Hyyppä, J., Saari, H., Pölönen, I., Imai, N.N., Tommaselli, A.M.G., 2017. Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging. Remote Sens. 9 (3), 185. https://doi.org/10.3390/rs9030185.

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. Appl. Geomat. 6 (1), 1–15.

Näsi, R., Honkavaara, E., Lyytikäinen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., Viljanen, N., Kantola, T., Tanhuanpää, T., Holopainen, M., 2015. Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. Remote Sens. 7 (11), 15467–15493. https://doi.org/10.3390/rs71115467.

NLS, 2016. National Land Survey of Finland, National Land Survey open data licence—version 1.0—1 May 2012, http://www.maanmittauslaitos.fi/en/NLS_open_data_licence_version1_20120501. (accessed 24 June, 2016).

Oliveira, R.A., Tommaselli, A.M., Honkavaara, E., 2016. Geometric calibration of a hyperspectral frame camera. Photogram. Rec. 31 (155), 325–347. https://doi.org/10.1111/phor.12153.

Rosnell, T., Honkavaara, E., 2012. Point cloud generation from aerial image data acquired by a quadrocopter type micro unmanned aerial vehicle and a digital still camera. Sensors 12 (12), 453–480. https://doi.org/10.3390/s120100453.

Saari, H., Pellikka, I., Pesonen, L., Tuominen, S., Heikkilä, J., Holmlund, C., Mäkynen, J., Ojala, K., Antila, T., 2011. Unmanned Aerial Vehicle (UAV) operated spectral camera system for forest and agriculture applications. In: SPIE Proc. 8174, 8174OH. doi: 10.1117/12.897585.

Saari, H., Pölönen, I., Salo, H., Honkavaara, E., Hakala, T., Holmlund, C., Mäkynen, J., Mannila, R., Antila, T., Akujärvi, A., 2013. Miniaturized hyperspectral imager calibration and UAV flight campaigns. In: SPIE Proc. 8889, 88891O. doi: 10.1117/12.2028972.

Schaepman, M., 2009. Imaging spectrometers. In: Warner, T.A., Nellis, M.D., Foody, G.M. (Eds.), The SAGE Handbook of Remote Sensing. SAGE Publishers Ltd, pp. 166–178. https://doi.org/10.4135/9780857021052.n12.

Senop, 2017. http://senop.fi/en/optronics-hyperspectral (accessed 26 August, 2017).

Specim, 2017. Specim, Spectral Imaging Ltd. http://www.specim.fi/hyperspectral-remote-sensing (accessed 26 August, 2017).

Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., Kooistra, L., 2014. A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. Remote Sens. 6, 11013–11030. https://doi.org/10.3390/rs61111013.

Tommaselli, A.M., Oliveira, R.A., Nagai, L.Y., Imai, N.N., Miyoshi, G.T., Honkavaara, E. and Hakala, T., 2015. Assessment of bands coregistration of a light-weight spectral frame camera for UAV. Proc. GeoUAV - ISPRS Geospatial Week, La Grande Motte, France, 192. https://geouav.teledetection.fr/papers/GEOSPATIAL_WEEK_2015_192.pdf.

Tommaselli, A.M.G., Berveglieri, A., Oliveira, R.A., Nagai, L.Y., Honkavaara, E., 2016. Orientation and calibration requirements for hyperspectral imaging using UAVs: a case study. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences 40 (3/W4), 109–115. doi: 10.5194/isprsarchives-XL-3-W4-109-2016.

Torres-Sanchez, J., Lopez-Granados, F., De Castro, A.I., Pena-Barragan, J.M., 2013. Configuration and specifications of an Unmanned Aerial Vehicle (UAV) for early site specific weed management. PLoS One 8 (3), e58210. https://doi.org/10.1371/journal.pone.0058210.

Turner, D., Lucieer, A., Malenovský, Z., King, D.H., Robinson, S.A., 2014. Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over Antarctic moss beds. Remote Sens. 6 (5), 4003–4024. https://doi.org/10.3390/rs6054003.

Uto, K., Seki, H., Saito, G., Kosugi, Y., 2013. Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 6 (2), 851–860. https://doi.org/10.1109/JSTARS.2013.2250921.

Vakalopoulou, M., Karantzalos, K., 2014. Automatic descriptor-based co-registration of frame hyperspectral data. Remote Sens. 6 (4), 3409–3426. https://doi.org/10.3390/rs6043409.

Vane, G., Green, R.O., Chrien, T.G., Enmark, H.T., Hansen, E.G., Porter, W.M., 1993. The airborne visible/infrared imaging Spectrometer (AVIRIS). Remote Sens. Environ. 44 (2–3), 127–143. https://doi.org/10.1016/0034-4257(93)90012-M.

Wu, C., 2013. Towards linear-time incremental structure from motion. In: International Conference on 3D Vision 3DV 2013, Seattle, WA, pp. 127–134. doi: 10.1109/3DV.2013.25.

Wytham Woods, 2017. Wytham Woods the Natural Place for Science http://www.wytham.ox.ac.uk/, (accessed 26 August, 2017).

Yang, Q., Ji, P., Li, D., Yao, S., Zhang, M., 2014. Fast stereo matching using adaptive guided filtering. Image Vis. Comput. 32 (3), 202–211. https://doi.org/10.1016/j.imavis.2014.01.001.

Zarco-Tejada, P.J., Gonzalez-Dugo, V., Berni, J.A.J., 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. Remote Sens. Environ. 117, 322–337. https://doi.org/10.1016/j.rse.2011.10.007.

Zitová, B., Flusser, J., 2003. Image registration methods: a survey. Image Vis. Comput. 21 (11), 977–1000. https://doi.org/10.1016/S0262-8856(03)00137-9.