ASSESSMENT OF BANDS COREGISTRATION OF A LIGHT-WEIGHT SPECTRAL FRAME CAMERA FOR UAV

Antonio M. G. Tommaselli [1], Raquel A. Oliveira [1], Larissa Y. Nagai [1], Nilton N. Imai[1], Gabriela T. Miyoshi [1], Eija Honkavaara [2], Teemu Hakala [2]

(1) Univ Estadual Paulista, Unesp, Brazil *(2) Finnish Geospatial Research Institute, FGI, Finland*

Highlights: A light-weight spectral frame camera is being integrated to an UAV for tropical forest monitoring. The images are not acquired instantaneously and the bands are miss-registered due to sensors miss-alignment and camera motion. The results of coregistration between bands are presented and discussed for two case studies with different camera models and environments.

Key words: *Light-weight Spectral Camera; coregistration; 2D mapping functions*

Introduction

Imaging acquisition based on unmanned airborne vehicles (UAVs) is becoming a mature activity and fostering several application fields because of the flexibility and favourable cost-benefit ratio, especially when high temporal resolution is required [1]. However, several problems cannot be properly approached with visible bands of conventional optical sensors and, in such case, multi or hyper spectral sensors are required.

Hyperspectral sensors usually scan the scene with a pushbroom geometry, covering instantaneously several spectral bands of a scanning row forming images which are combined in a data cube. However, due to the line scanning principle, each image row is captured from a different position and attitude, and these orientation data have to be acquired during the flight using dual frequency GNSS receivers and integrated Inertial Measurement Units. Albeit these equipment are becoming more affordable, the costs of units which provide the required accuracy are sometimes higher that the UAV itself. To avoid the need for higher grade IMUs, hyperspectral frame cameras can be used and the required EOPs (Exterior Orientation Parameters) can be indirectly computed via bundle adjustment. A light-weight hyperspectral frame camera was developed by VTT Technical Research Center of Finland [6] which is suitable to be carried by an UAV and fits the requirements of several applications [3][4]. This light-weight spectral camera (< 700 g) is based on a Fabry-Perot interferometer (FPI) inserted between the sensor and the lens. The sensor spectral sensitivity is a function of the interferometer air gap and acquisition of different wavelengths is performed by changing this gap, being possible to reconstruct the spectrum for each pixel in the image. The same principle is being used by Rikola Ltd [5] to build up their cameras, one of them used in this research.

However, when the platform carrying this frame camera is moving, the spectral bands of the data cube do not overlap perfectly because images of the same data cube are acquired from different positions and orientations (Fig. 1). Besides some recent camera models have two sensors and their images are also not perfectly aligned. Due to these camera features, a registration process between images bands is required. The effect of platforms movement in the band registration, with different camera versions, can be seen in Figures 1.a and 1.b, the former being acquired by an UAV in Finland and the later in a terrestrial calibration field in Brazil. This effect is highly dependent on the platform speed, distance to the object and relief displacement.

Image registration aims at mapping pixels between images acquired under different conditions and requires a mapping function which parameters can be estimated either from the parameters of each image or from corresponding points in both images. After the parameters of the mapping function have been estimated, one image can be resampled with respect to the reference one. The main steps of image registration can be summarized as [1][2]: (1) Interest features (points, corners, edges) extraction; (2) Features matching; (3) Estimation of the mapping function parameters; (4) Image resampling.

Mapping functions usually are geometric transformations that can be either local or global. Global models are applied over the whole image while the local models divide the images in several regions for which different sets of parameters are estimated [2]. Examples of linear geometric transformations are: rigid body, Helmert and affine; projective transformation is non-linear. Besides these 2D transformations, polynomials equations can also be used and are common for registration of satellite images. The estimation of the parameters is usually done by Least Squares method from a set of overdetermined equations, established from corresponding points.

These procedures are suitable when the displacements between images are small and for flat terrain. Large base length and relief variations will introduce parallaxes in the images which cannot be corrected by a global mapping function. In such cases, orthorectification may be one strategy to produce perfectly registered bands, but at the cost of resampling the images and with greater algorithm complexity, requiring bundle adjustment and

generation of a digital surface model (DSM). Another major issue in the image registration task is the search for distinguishable features for matching. Depending on the terrain coverage, this process can be troublesome, mainly when homogeneous areas or repetitive patterns occur. The differences in spectral responses of some natural targets and shadowing also hinder the matching process, being necessary to use feature based techniques that should be invariant to gray levels differences between different bands. The well-known method based on correlation will work only for neighbor bands and for some targets. Low SNR is also common in hyperspectral cameras, due to the narrow bands, and this also imposes some problems in feature extraction and matching.

There are few papers on the registration between bands of this type of camera. Vakalopoulou and Karantzalos [7] used SURF and SIFT descriptors for matching and used group of bands, to avoid co-registering bands that are significantly displaced. Honkavaara et al. [4] also used the strategy of selecting some reference bands and matching the rest of the bands to those.

The aim of this study is to assess the quality of the co-registration between images from different spectral bands of the same data cube, using different transformations: 2D (Helmert, affine, projective and polynomial second order) using two camera models and with different scenarios (Fig. 1).

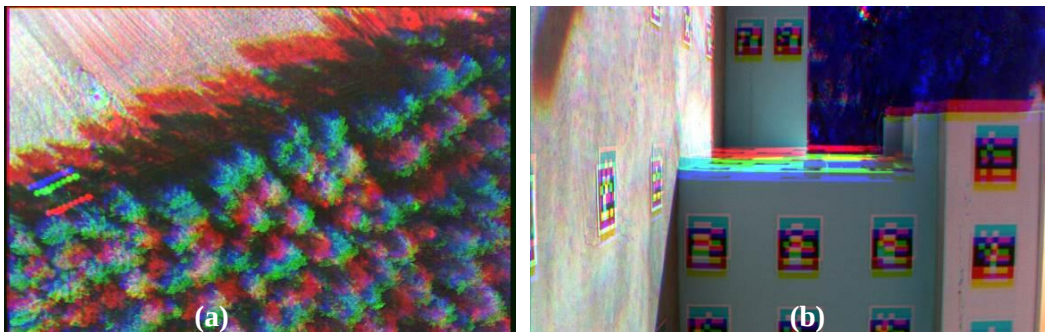

Figure 1: Band miss-registration due to platform movement: (a) spectral aerial image collected by an UAV and: (b) terrestrial images of a calibration field with coded targets, collected with a moving platform simulating the same apparent speed of an UAV.

Experiments

Two data sets were used in this work from different cameras and configurations.

The first data cube was acquired by a FPI camera, prototype 2012b, in **Evo, Finland**, with a GSD of 15 cm, carried by an octocopter UAV helicopter from FGI [4]. This camera model has the FPI, one single CMOS sensor with Bayer filter and the infrared filter was removed. Changing air gaps of the FPI modifies the narrow wavelength band produced [6]. Previously defined air gaps are applied to reconstruct the spectrum for each pixel in the image [4]. Acquisition of a single cube with 24 air gaps, takes a total of 1,800 ms. The images used in this case study have 1024x648 pixels and 27 bands, but only 16 bands were selected and only one band was used as reference (band 6: λ = 526.3, Table 1, which is the approximately the average time of cube acquisition). Considering a flight speed of 5 m/s, the displacement between the camera centers from the reference band 6 to band 13 was larger than 4 m. Interest points (24) were extracted by Förstner operator with further interactive inspection and editing to ensure the same points in all bands for the sake of comparison. The parameters of the geometric transformations were estimated and the results were assessed with independent 9 check points.

Table 1: Selected bands and their differences in position (ds) and time of acquisition (dt) with respect to a reference band (6).

Band n°	**13**	**14**	**15**	**16**	**17**	**3**	**4**	**24**	**6**	**8**	**11**	**7**	**10**	**23**	**5**	**18**
λ (nm)	571	593.8	608.4	614.7	628	593.8	516.2	757.8	526.3	538.9	551.6	527	549.6	726.7	521.2	699.5
dt to RefBand (s)	-0.83	-0.75	-0.68	-0.6	-0.53	-0.15	-0.08	-0.08	0	0.075	0.15	0.3	0.45	0.45	0.6	0.75
ds to RefBand (m)	-4.12	-3.75	-3.37	-3	-2.62	-0.75	-0.37	-0.37	0	0.375	0.75	1.5	2.25	2.25	3	3.75

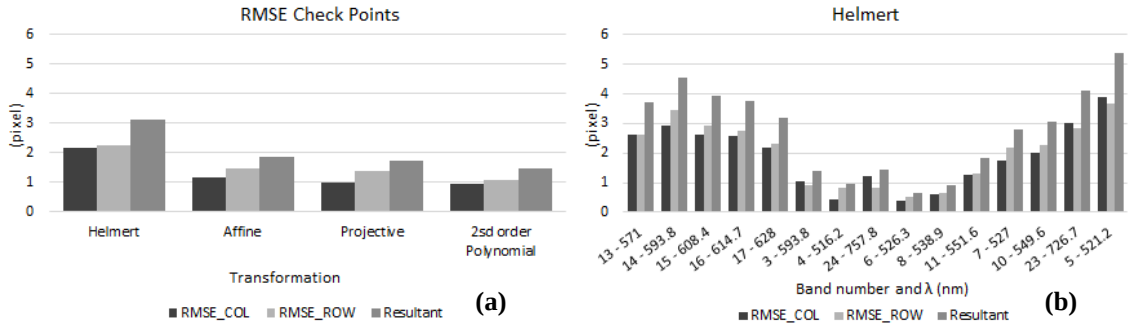

Figure 2: (a) RMSE in check points after estimation of transformation parameters for all band pairs; (b) RMSE in check points for the Helmert transformation.

The parameters of four selected mapping functions (Helmert, Affine, Projective, second order Polynomial) were estimated for each band pair, from 24 corresponding points. Fig. 2.a presents the root mean square error (RMSE) of the discrepancies in the 24 control points and also for 9 checkpoints. Results presented in Fig. 2.a show that the best results were achieved with polynomial functions. The results for each pair band showed that the further is the band with respect to the reference band the higher are the RMSE, which can be seen in Fig. 2.b, for Helmert transformation. This can be explained by the local image displacements caused by relief variations (parallax) that cannot be corrected by global mapping functions.

The second set of data cubes was acquired by a Rikola DT-0014 camera [5] in a terrestrial calibration test field at Unesp, in Brazil. This camera also uses the FPI but with two CMOS sensors and without Bayer filter. This improves the image quality but introduces a further problem with the sensors geometric alignment. The image of datacubes have 1023x648 pixels, with 25 bands but only 5 bands were selected for the this registration assessment and one band was used as reference (band n. 13, Table 2). The camera was mounted over a horizontal supporting table (Fig. 3) and installed over a pallet carrier. The distance from the camera to the wall with coded targets was 6.9 m and the camera focal length 8.7 mm. The platform was moved with two different speeds (0.16 m/s and 0.55 m/s) in order to produce apparent speeds comparable to flights with UAV speeds of 3.4 m/s and 11 m/s, respectively, for a flying height of 150 m. Two image strips were collected and, from this data, 3 data cubes of each strip were selected to have samples of different situations: (1) camera static; (2) moving camera with flat terrain; (3) moving camera with high terrain variations. Table 2 presents the bands selected for comparison: Bands 1 and 7 are acquired by sensor 2 while bands 13, 19 and 25 are acquired by sensor 1. Interest points were automatically extracted and some additional points were interactively measured. Using pairs of corresponding points (control points) the image of four bands were registered with the reference band (13) using affine transformation and 2nd order polynomial. The experiments were performed with: (a) static camera; (b) low speed (0.16 m/s) and higher speed (0.55 m/s). The configuration of control points were variable and three scenarios were considered (see Fig. 1.b): (a) control points over a flat area; (b) control points with depth variations of 1m, 10% of the camera distance (c) depth variation of 3 m, 30% of the camera distance.

Table 2: Selected bands of RIKOLA DT camera and their differences in position (ds) and time of acquisition (dt) with respect to a reference band (13).

Band n°		**1**	**7**	**(Ref) 13**	**19**	**25**
λ (nm)		506.07	592.78	669.96	729.56	819.74
dt to RefBand (s)		-0.4	-0.17	0	0.17	0.4
ds to RefBand v = 0.55 m/s	cm	- 22.00	-9.35	0.00	9.35	22.00
	pixels	-54.24	-23.05	0.00	23.05	54.24
ds to RefBand v= 0.16 m/s	cm	-6.40	-2.72	0.00	2.72	6.40
	pixels	-16.13	-6.86	0.00	6.86	16.13

Figure 3: The RIKOLA DT camera mounted on a terrestrial platform.

The RMSE of the residuals in the control points are presented in Tables 3, 4 and 5. Table 3 presents the results with: (a) the camera static and for control points in a flat area and; (b) with variations around 10 % of the camera distance. It can be seen that the differences between the transformations are smaller when the bands are taken by the same sensor (13, 19 and 25). In these cases the results achieved for flat area are similar, when using the affine and polynomial transformations. When registering bands of sensor 2 (1 and 7) to band 13 (sensor 1) the polynomial function presents better results, probably because it can absorb the miss-alignment between the two sensors and other distortions. The magnitudes of residuals are higher when the control points do not fit in a flat surface, but the residuals are still subpixel.

Table 3: RMSE (in pixels) in control points used to estimate transformation parameters for band pairs with **static** camera: (a) refers to control points in a flat area and (b) controls points with 1 m of depth variation.

	(a) flat area				(b) 10 % of depth variation			
Transformation \Bands pairs	1-13	7-13	19-13	25-13	1-13	7-13	19-13	25-13
Affine	0.45	0.43	0.20	0.30	0.74	0.72	0.33	0.38
2nd order Polynomial	0.36	0.36	0.19	0.28	0.64	0.60	0.33	0.34

Table 4: RMSE (in pixels) in control points (CP) used to estimate transformation parameters for band pairs with **camera moving at 0.16 m/s**: (a) refers to CPs in a flat area and (b) CPs with 1 m of depth variation.

	(a) flat area				(b) 10 % of depth variation			
Transformation \Bands pairs	1-13	7-13	19-13	25-13	1-13	7-13	19-13	25-13
Affine	0.54	0.57	0.33	0.37	1.21	0.98	0.70	1.63
2nd order Polynomial	0.38	0.36	0.27	0.29	0.76	0.72	0.48	0.90

Table 4 presents the results with the camera moving at 0.16 m/s (corresponds to an UAV speed of 3.4 m/s) and for: (a) control points in a flat area and; (b) with variations around 10 % of the camera distance. The RMSE when registering between different sensors are again higher. Considering that the camera is moving, the parallax effect can be seen in the RMSE on CPs with depth variations. For instance, when using affine transformation, the registration of bands taken in the beginning and end of the data cube (1 and 25) produces residuals with RMSE greater than 1 pixel. Polynomial model reduces theses residuals, but they are still near 1 pixel.

Table 5: RMSE (in pixels) in control points used to estimate transformation parameters for band pairs with **camera moving at 0.55 m/s**: (a) refers to CPs in a flat area and (b) controls points with 3 m of depth variation.

	(a) flat area				(b) 30 % of depth variation			
Transformation \Bands pairs	1-13	7-13	19-13	25-13	1-13	7-13	19-13	25-13
Affine	0.43	0.72	0.41	0.43	11.00	5.60	6.18	12.50
2nd order Polynomial	0.19	0.22	0.30	0.19	3.60	2.08	1.88	3.33

Table 5 presents the results with the camera moving at higher speed: 0.55 m/s (corresponds to an UAV speed of 11 m/s) and for: (a) control points in a flat area and; (b) with variations around 30 % of the camera distance (3 m). It becomes clear from the results, that for flat areas the images can be registered but for areas with significant depth variations, the parallax effect caused by the camera displacement during bands acquisition cannot be corrected with 2D transformations. The RMSE presented in Table 5, sixth to ninth columns, are higher than one pixel and this means that pixels will not be correspondent in the warped image.

Several problems can be foreseen in the registration of bands taken by this kind of spectral frame cameras in forest areas: the detection of corresponding points is difficult, due to the different spectral responses of the targets, and because of the target characteristics such as homogeneity and shadows. The camera displacements during band acquisition cause significant parallax effects, depending on the relief variations. In some cases, it will be unreliable to register bands with 2D transformations. In these cases, it can be necessary to reduce the flight speed depending on the object height variations or to adjust the GSD output to meet the registration accuracy. Another alternative to be investigated is the use of pixel to pixel discrete mapping based on digital surface models and relative orientation. The camera is now integrated to an octocopter and experiments are being performed with real aerial images for which the technique of discrete mapping will be tested.

Acknowledgements

The authors would like to acknowledge the support of FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) through a research grant (2013/50426-4) and a PhD Scholarship (nº 2013/17787-3) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) through a research grant (307554/2014-7) and a MSc Scholarship (nº 130871/2014-1).

References

[1] Brown, L. G. (1992). A survey of image registration techniques. ACM computing surveys (CSUR) 24 (4), 325-376.

[2] Zitová, B., Flusser, J. (2003). Image registration methods: a survey. Imaging and Vision Compunting, 21 (11) 977-1000.

[3] Colomina, I., Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 79-97.

[4] Honkavaara, E., Saari, H., Kaivosoja, J., Polonen, I., Hakala, T., Litkey, P., Makynen, J., Pesonen, L. (2013). Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. Remote Sensing. 5 (10), 5006–5039.

[5] RIKOLA (2014) http://www.rikola.fi/site/products/hyperspectral-camera/.

[6] Saari, H.; Aallos, V. V.; Akujärvi, A.; Antila, T.; Holmlund, C.; Kantojärvi, U.; Ollila, J. (2009). Novel miniaturized hyperspectral sensor for UAV and space applications. In SPIE Europe Remote Sensing. International Society for Optics and Photonics, 74741M-74741M-12.

[7] Vakalopoulou, M., Karantzalos, K. (2014). Automatic Descriptor-Based Co-Registration of Frame Hyperspectral Data. Remote Sensing, 6 (4), 3409–3426.