

Distortion Invariant Optical Calibration for Lightweight Airborne Imaging Spectrometry

Julien Burkhard¹, Jesse Lahaye¹, Laurent V. Jospin¹, Jan Skaloud¹

¹ Environmental Sensing Observatory (ESO), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland – julien.burkhard@alumni.epfl.ch, (jesse.lahaye, laurent.jospin, jan.skaloud)@epfl.ch

Keywords: Unmanned Aerial Vehicles, Pushbroom Sensor, Airborne Imaging Spectrometer, Bayesian Modeling

Abstract

Hyperspectral cameras have recently been miniaturized for operation on lightweight airborne platforms such as UAV or small aircraft. Unlike frame cameras (RGB or Multispectral), many hyperspectral sensors use a linear array or ‘push-broom’ scanning design. This design makes image rectification and the calibration of intrinsic and extrinsic camera parameters significantly more challenging. Standard methods employed to address such tasks rely on precise GPS/INS estimates of the airborne platform trajectory and a detailed terrain model. However, inaccuracies in the trajectory or surface model can introduce systematic errors and complicate geometric modeling, which ultimately degrade the quality of the image rectification. To overcome these challenges, we propose a method for tie point extraction and camera calibration for ‘push-broom’, hyperspectral sensors using only raw spectral imagery and a raw, possibly low quality, GPS/INS trajectory. We demonstrate that our approach allows for the automatic calibration of airborne systems with push-broom, hyperspectral cameras, outperforms other state-of-the-art automatic rectification methods and reaches an accuracy on par with manual calibration methods.

1. Introduction

Until recently, hyperspectral sensors (HS) have not been practical to use as payloads for light airborne platforms (e.g. UAVs) (Nex et al., 2022), unlike frame cameras. Lightweight HS offer many advantages over larger, aircraft based sensors. They offer higher image resolution, greater deployment flexibility and lower cost. These aspects make them highly useful for a wide range of applications, such as snow-cover characterization (Ding et al., 2022), crop and forest monitoring (Asner, 1998), and more. To maximize the achievable spatial and spectral resolution, most light-weight Airborne Imaging Spectrometers (AIS) operate as push-broom cameras (Nex et al., 2022). This makes georectifying their data challenging, as the solution is heavily reliant on the GPS/INS solution associated with the imagery. Orientation determination with inertial measurements remains an especially challenging task for low-end inertial sensors (Sharma et al., 2024). There are two major types of distortions which affect the output of push-broom sensors and lead to the aforementioned georeferencing inaccuracy (Fang et al., 2017): high frequency perturbations caused by non-stabilized platforms on multirotor drones and low frequency attitude drift caused by GPS/INS imprecision. In general, roll motion causes horizontal shift between successive lines in the image, while pitch and speed variation cause along-track deformation (see Figure 1).

Imprecision of the georeferenced data becomes particularly problematic when multiple data modalities are to be considered together, e.g., point cloud and HS images for surface characterization (Liu et al., 2017), or when a time series of the data must be considered, e.g., for snow cover monitoring (Dozier and Painter, 2004). Another challenge arises when the GPS/INS sensor is attached to the airborne platform rather than to the hyperspectral sensor. In this configuration, the boresight and lever arm between the camera projective center and the IMU center, must be frequently re-calibrated. In fact, most hyperspectral cameras integrate an internal GPS/INS sensor to solve this issue. This makes the payload heavier, requiring larger, more

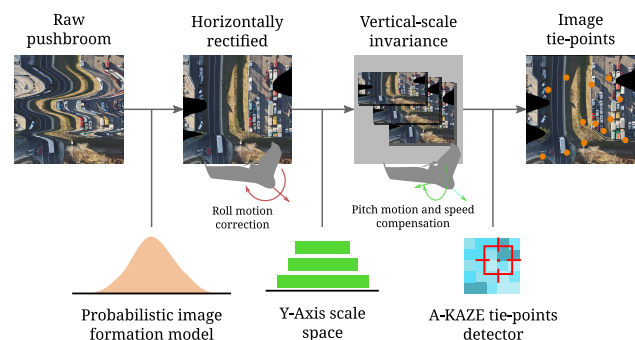


Figure 1. Overview of our proposed process for tie-point extraction in push-broom hyperspectral imagery

expensive drones and limits the capability of the camera to be operated jointly with other payloads. In addition, this often only reduces the amplitude of the boresight, leaving precise calibration of the sensor to the user (Sankararao et al., 2020).

Computer vision methods such as bundle adjustment can be used to mitigate errors in the trajectory and mounting offsets and rectify the image orientation (Barbier et al., 2016). However, the processing methodologies for push-broom imagery remain much more complex than the equivalent process for frame imagery, widely available in commercial software suites such as Agisoft Metashape (Agisoft, 2024) or Pix4D (SA, 2024). There are two main reasons for this additional complexity. First, the motion of the airborne platform can cause misalignments between pushbroom scan line, complicating the automatic extraction of accurate tie points from the data. Second, in push-broom imagery, each line has a different pose, which dramatically increases the number of parameters considered in the bundle adjustment, and introduces certain limits on the geometric constraints that can be used when filtering the data (e.g., epipolar constraints (Zhang, 2021)). These limitations prevent the use of algorithms commonly used for frame images, such as the 8-point algorithm (Hartley, 1997), to align push-broom imagery.



Figure 2. Pika-L HS camera mounted on a DJI M350

In this work, we contribute a new method for automatic tie point extraction from push-broom, HS images without the requirement of additional data (i.e., terrain model or GPS/INS data):

1. We propose a probabilistic model of push-broom image formation to rectify high-frequency distortion caused mainly by roll motion in the image data. This approach yields subpixel accurate estimates of the shift between image lines, and does so without access to a terrain model or GPS/INS data.
2. We propose a y-scale invariant matching-scheme to resolve distortions caused by pitch motions and speed differences when matching push-broom images. We show that this significantly increases the number of matches and inlier ratio.
3. We then show a practical use of the proposed methods by evaluating the quality of boresight calibration of AIS cameras using the proposed visual matches. Tie points generated with our algorithm, coupled with a robust algorithm for boresight estimation, provide an orientation accuracy of approximately 0.2° , employing a low cost IMU (SBG Ellipse-N), compared to 0.3° for state-of-the-art tie point generation methods.

Next, in section 2 we review related works in the literature. In section 3 we review the proposed model. In section 4 we present the experimental methodology we used to evaluate our model, as well as the results we obtained. We conclude in section 5.

2. Related works

Processing pipelines for push-broom AIS imagery usually rely only on the GPS/INS solution to georectify image lines from the camera (Sankararao et al., 2020, Nex et al., 2022). To limit the impact of high frequency vibrations, stabilization gimbals can be used (Hueni et al., 2025, Turner et al., 2017) but they pose other issues (i.e., additional weight and size, time varying boresight and lever arm between the GPS/INS sensor and the camera).

Historically, strong straight features (e.g. roads) had to be manually identified to correct for distortion caused by aircraft motion (mostly roll motion) (Jensen et al., 2008), but such methods are not feasible in the absence of a straight, linear pattern/object in the image, and are complex to automate. In addition to linear features, Ground Control Points (GCPs), marked with recognizable patterns that can be identified in the images and for which spatial coordinates are measured on the terrain, can also be used as constraints (Ryan R. Jensen and Jensen,

2011). Alternatively to GCPs, which require an intervention in the terrain, image correspondences or tie points can be extracted from overlapping image regions. Tie points have been widely used for HS sensor calibration, either extracted from overlapping HS cubes (Berveglieri and Tommaselli, 2019), or between HS cubes and frame images (Barbieux et al., 2016). However, the reliable extraction of tie points can be challenging in the presence of distortion in HS imagery. This presents a conundrum when such tie points are required to rectify these deformations in the first place. In such cases, the image can be pre-rectified using only the raw solution from the GPS/INS sensor. However, artifacts can be generated in the image during the rectification process, if there is no prior boresight calibration, the GPS solution is noisy, or there is a time offset between the cameras and the GPS/INS sensor (Habib et al., 2018).

It has been shown that the cross-correlation of pixels between successive scan lines can be used to correct the horizontal displacement caused by aircraft roll motion (Fang et al., 2017). This makes it possible to extract tie-points from the partially rectified hyperspectral cubes, without any form of direct georeferencing of the data or boresight calibration. We propose a similar approach for horizontal displacement correction, but we replace the cross-correlation with a Bayesian probabilistic model (Section 3.2), which has multiple benefits over the state of the art. The Bayesian model better characterizes the stochastic properties of the image, enabling sub-pixel accuracy in predicting the horizontal shift between consecutive image scan lines. Additionally, it incorporates a prior which regularizes the solution for improved stability.

For tie point extraction in HS imagery, spatial/spectral variants of popular descriptors such as SIFT have been proposed (Yu et al., 2021), but in practice, it has been observed that traditional descriptors yield reliable results without the need to be adapted (Berveglieri and Tommaselli, 2019). In this work, we used the standard A-KAZE matching pipeline (Alcantarilla and Solutions, 2011) to detect tie points after image rectification.

3. Methodology

The proposed method aims to obtain tie points for HS, push-broom imagery in the presence of either a non-stabilized platform, inaccurate GPS/INS measurements, time tagging offset w.r.t GPS time and/or unknown boresight between the GPS/INS system and the camera. We split the main problem into three: 1) horizontal shift correction due to roll motion, 2) removing the ambiguity in vertical scaling due to pitch motion and variable speed and finally 3) tie point extraction.

Section 3.1 presents the dataset used in our experiments. Section 3.2 presents our method for horizontal shift correction, based on a probabilistic model. Section 3.3 presents our approach for dealing with the ambiguity in vertical scaling of the raw AIS data. Finally, Section 3.4 presents our approach for tie-point extraction.

3.1 Datasets

The dataset for this work was acquired in the region of Delemont, in Switzerland. A Resonon Pika-L HS camera was used to acquire the data. With a total payload weight of approximately 1kg, this camera is ideal to be flown on certain UAVs, such as the DJI M350, see Figure 2. For this experiment, the camera was mounted on a helicopter, alongside a reference navigation

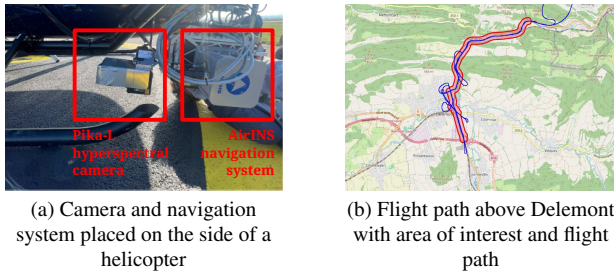


Figure 3. Dataset acquisition in Delemont (CH), map background © OpenStreetMap contributors

grade GPS/INS system comprised of a Javad GNSS receiver and iXblue AirINS inertial platform, see Figure 3. The Pika-L comes bundled with a GPS-INS system, SBG Ellipse-N, which is representative of the quality one could expect by flying the camera on a UAV, while the output of the AirINS was used as a ground truth reference. The flight lasted approximately 1 hour and was conducted at an altitude of 150 m above ground level, with a flight speed of 14 m/s. The Pika-L, with a field of view of 36.5° and 900 spatial pixels per scan-line, was operated at a sampling frequency of 200 Hz. The resulting ground sampling distance of the HS data was approximately 10 cm/px.

A hardware output pulse triggered for each exposure by the Pika-L was used to time-tag scan-lines with absolute GPS time. Within this configuration, more pulses sent from the Pika-L were registered by the GNSS receiver than scan-lines were recorded. This is likely due to hardware limitations and sensitivity to noise of the acquisition setup. Nonetheless, since our method for horizontal shift rectification, presented in Section 3.2, does not require GPS/INS data, but the horizontal shift is correlated with roll, we were able to recover the correct time offsets between the captured HS scan-lines and the absolute GPS-time of the pulses.

3.2 Horizontal shifts and image formation model

At the core of our approach, we consider that the value of two successive lines in the push-broom images, represented here as a vector I , follow approximately a normal distribution whose covariance is a function of the relative pose $\Delta\Gamma$ between two lines:

$$I \sim \mathcal{N}(\mu, \Sigma(\Delta\Gamma)). \quad (1)$$

This derives from the assumption that the image generation can be approximated by a Gaussian Process, an assumption, while not perfect, has been shown to be sufficiently accurate in practice when correct covariance functions are used (He and Siu, 2011). If we have a prior on how different pixels are correlated based on the relative poses of the camera between the two lines, then observing I yields insight about the relative orientation of the camera between two lines by inverting the conditional probability distribution $p(I|\Delta\Gamma)$ to obtain $p(\Delta\Gamma|I)$. A straightforward way of doing so is to use the Bayesian formalism and probabilistic graphical models (Buntine, 1994).

In practice, a model based on the full 6-DOF relative pose $\Delta\Gamma$ will be ambiguous, as the posterior $p(\Delta\Gamma|I)$ can be a complex multi-modal distribution in which effects such as change in pitch or forward motion can be highly correlated. To remove

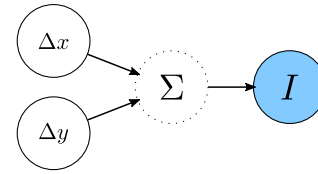


Figure 4. Probabilistic graphical model of our approach

the ambiguity, we only consider the resulting horizontal (cross-path) shift Δx and vertical (along-path) shift Δy , as these are sufficient to pre-rectify the push-broom data and find tie points. With the Pika-L optics, the relative error of approximating the effect of roll motion with horizontal shifts is on average 0.03 pixels, which is lower than our target estimation accuracy of 0.1px.

We build our probabilistic graphical model, see (Buntine, 1994) for more details, as shown in Figure 4. We set the prior such that $p(\Delta x)$ is a normal distribution with mean equal to 0 px and standard deviation equal to 0.5 px and $p(\Delta y)$ is an exponential distribution with rate parameter equal to 1. The values of the standard deviation and rate were chosen empirically by analyzing the raw IMU data.

The likelihood model is given, following Equation 1, by:

$$p(I|\Delta x, \Delta y) = \mathcal{N}(\mu, \Sigma(\Delta x, \Delta y)), \quad (2)$$

where μ is the mean of the pixel value intensity, and the function $\Sigma(\Delta x, \Delta y)$ is computed using a Matérn covariance kernel (Genton, 2002). With this model, the covariance between two pixels separated by a horizontal distance d is given by:

$$\sigma^2 \left(1 + \frac{\sqrt{3}\sqrt{(\Delta x + d)^2 + \Delta y^2}}{l} \right) e^{\left(-\frac{\sqrt{3}\sqrt{(\Delta x + d)^2 + \Delta y^2}}{l} \right)}, \quad (3)$$

with σ^2 the variance of the image pixel intensity and l a length factor that can be calibrated by evaluating the covariance of pixels on the same row (thus without distortion). We set $\Delta x = \Delta y = 0$ for pixels on the same row. The Matérn covariance function ensures that the covariance between two pixels decreases effectively with distance, while ensuring that $\Sigma(\Delta x, \Delta y)$ remains positive definite for all values of Δx and Δy , which an arbitrary covariance function cannot guarantee.

In the end, we select the estimator for Δx and Δy to be:

$$\Delta \hat{x}, \Delta \hat{y} = \arg \max_{\Delta x, \Delta y} p(I|\Delta x, \Delta y)p(\Delta x)p(\Delta y). \quad (4)$$

A closed form expression for $\log(p(I|\Delta x, \Delta y)p(\Delta x)p(\Delta y))$ and its derivative exists, however solving for the minima is not feasible, thus we rely on the l-BGFS (Liu and Nocedal, 1989) numerical solver to compute $\Delta \hat{x}$ and $\Delta \hat{y}$.

3.3 Vertical scale invariance

It was observed that $\Delta \hat{x}$ yields a good estimate of the horizontal shift and is very correlated with the roll motion of the airborne

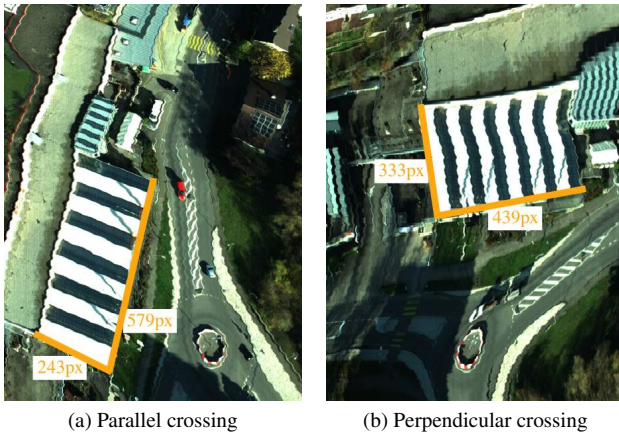


Figure 5. The effect of y-scaling, where the aspect ratio of the image changes when the airborne platform flies parallel (a) or perpendicular (b) to the road.

platform. On the other hand, $\Delta\hat{y}$ is usually less predictable and yields very little information about the local distortion along the direction of motion. This is because when sharp transitions in the landscape occur (e.g., the plane crosses the edge of a building along the axis of travel), the local correlation hypothesis is no longer valid. In that case, $\Delta\hat{x}$ is only marginally affected, as the transition occurs for only a single pixel in the across-track direction, while, since multiple pixels can be part of the transition in the along-track direction, $\Delta\hat{y}$ is strongly affected. For our intended application, we need a different method to treat vertical distortion in the HS data.

This is because, for crossing flight lines, i.e., when tie points are the most useful for calibration purposes, the variable speed of the aircraft can lead to distortion in the push-broom pixels along the direction of travel, see Figure 5. These distortions, in turn, complicate the process of matching keypoints as they act in opposite directions in the two crossing flight lines.

To deal with possible distortions in the y direction, we modified the A-KAZE scale invariant feature pyramid, such that the scale in the x and y directions are now explored independently. This anisotropic scale pyramid trades speed and complexity of the algorithm for invariance against y scaling. Nonetheless, in aerial imaging, the aircraft usually flies at constant height above the ground, so very few levels in the x direction need to be considered. Generally a single level and rarely more than 2 different levels are sufficient. In our study we only considered a single scale level in the x direction, as the whole feature pyramid is meant to deal with the y scaling of the data.

3.4 Tie point extraction and filtering

Once the image is rectified as described in Section 3.2, it becomes possible to use a traditional pipeline to extract matching tie points. We used the A-KAZE method (Alcantarilla and Solutions, 2011), which is known to yield fast and accurate results (Ordonez et al., 2018), and is open source. Initial testing on our data hinted that A-KAZE is a better alternative to SIFT as it yields more matches and a similar proportion of inliers. We also compared with ORB (Rublee et al., 2011), (Tareen and Saleem, 2018). While ORB yields more matches, more of them are outliers, which complicates downstream tasks.

Once the tie point matches are produced, a significant portion can still be incorrect (i.e. the algorithm matches two

pixels representing different elements in the scene). In the case of frame images, geometric constraints and the RANSAC algorithm (Lacey et al., 2000), commonly epipolar constraints, can be used to detect and remove outliers. In the case of uncalibrated pushbroom cameras, assuming the GPS/INS measurement is either unreliable or imprecise, this is not feasible. Instead, we consider two rectified chunks of the pushbroom HS data and assume that a generic homography exists between the points in chunk 1 and chunk 2. In other words, we assume the existence of a unique matrix H such that:

$$pt_{1,i} \propto Hpt_{2,i}, \quad (5)$$

where $pt_{1,i}$ and $pt_{2,i}$ are matching points in chunks 1, respectively 2, in homogeneous coordinates. The assumption here is that there exists a perspective transformation between the sets of inlier points in the two respective chunks.

The coefficients of H can be estimated with at least 4 points using Equation 5 and the direct linear transform algorithm (Abdel-Aziz et al., 2015). Based on this model, a RANSAC loop can be built. In the RANSAC loop, we consider a point to be an inlier if it lies at a distance of at most 60 pixels, which is very tolerant to account for the fact that our homography model is only an approximation.

3.5 Hyperspectral camera boresight calibration

The main parameter we want to calibrate for a HS camera is the boresight. Intrinsic parameters such as the focal length and lens distortion can be calibrated in advance, meaning that a strong prior can be attributed to them. The lever arm between a GPS/INS sensor and the camera can be measured on the platform before the flight with good accuracy (relative to the scale of the scene). On the other hand, the boresight, which is key for precise geo-referencing of the mapping data, cannot be measured. Tie points offer a good system for the boresight calibration from the mapping data itself, limiting the need for a separate calibration procedure before flight. Due to geometric constraints, it still imposes requirements on the flight path, as crossing lines are required to disambiguate the projected effects between the different axes of rotation.

To estimate the boresight, we minimize the residuals of epipolar constraints for each pair of tie points. We parameterized the boresight as an axis angle r , in which the direction indicates the axis of rotation, and the norm indicates the angle. This is in fact an element of $\mathfrak{so}(3)$, the Lie algebra associated with the special orthogonal group $SO(3)$. This means that the corresponding rotation matrix R is given by the matrix exponential $e^{[r]_{\times}}$, with $[r]_{\times}$ the matrix representing the cross product with r . The epipolar constraint for a pair of image points is written as:

$$\langle e^{[r]_{\times}} v_1 \times R_2^1 e^{[r]_{\times}} v_2, t_2^1 \rangle = 0, \quad (6)$$

with v_1 and v_2 being the ray direction in the camera frame at pose 1, respectively 2, R_2^1 being the relative rotation from pose 2 to pose 1 and t_2^1 being the translation between pose 1 and 2. R_2^1 and t_2^1 are taken from the GPS/INS output, and as such would be noisy, but the estimate of r is obtained with a very large number n hundred, if not thousands of tie points,

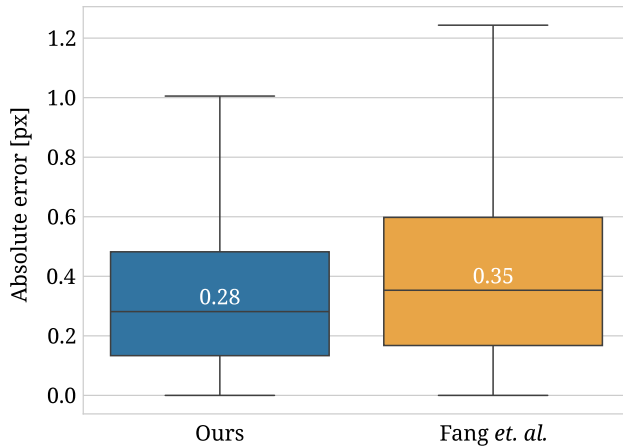


Figure 6. Horizontal shift error distribution (compared to Δx computed from ground truth GPS/INS)

lowering the noise level with a factor $\mathcal{O}(\sqrt{n})$ as long as no bias is present.

Equations 6 can be solved using a numerical solver like Gauss-Newton or Levenberg–Marquardt. We used Gauss-Newton, as the problem proved to be well behaved and converges to the solution without a need for additional regularization.

The Gauss-Newton algorithm solves an over-determined set of equations in the least square sense, meaning that a single outlier can have a large effect on the solution. To limit the impact of outliers, it can be useful to wrap the residuals of the equations in a non-linear kernel. A popular option is the Huber Loss, which is quadratic close to zero and linear further away:

$$H_{\delta}(x) = \begin{cases} \frac{1}{2}x^2 & \text{for } |x| \leq \delta, \\ \delta \cdot (|x| - \frac{1}{2}\delta) & \text{otherwise.} \end{cases} \quad (7)$$

with δ being a hyper-parameter of the loss. We selected $\delta = 1/4$.

4. Experiments

In this section, we evaluate the different methods described in Section 3. First, we evaluate our proposed probabilistic model for horizontal shift correction in Section 4.1. Then, in Section 4.2, we evaluate how the proposed horizontal shift correction and y-scale invariant approach impact the production of tie points, and the proportion of outliers filtered out. Finally, in Section 4.3, we measure the quality of boresight calibration as a function of the produced automated correspondences.

4.1 Horizontal shift correction

To evaluate the quality of the proposed horizontal-shift correction method, we compare its performance against the state of the art method by (Fang et al., 2017).

We use the reference solution from the navigation grade GPS/INS to estimate the ground truth horizontal displacement $\tilde{x}$. We compute the shift in pixels of the features on the ground, based on the digital surface model of the terrain. We then average that shift to get $\tilde{x}$. Note that only the regions within the mapping area were included. Regions outside of the mapping

areas highlighted in Figure 3, where the helicopter is turning, are excluded.

Overall, for the Delemont dataset, our method reaches an average accuracy, measured by the RMSE, of 0.85 pixels. Note that the RMSE is very strongly influenced by a small set of outliers. When the median is considered instead, the accuracy is 0.28 pixels. For comparison, the previous state of the art method by (Fang et al., 2017), based on the cross correlation of successive lines, reaches a RMSE 0.93 pixels and a median accuracy of 0.35 pixels. Figure 6 shows the distribution of error with the median, as well as the 5%, 25%, 75% and 95% quantiles.

Observations in the rectified images (Figure 7) show that there are two main reasons why our method reaches this accuracy:

First, since our model is continuous (able to detect sub-pixel shifts), and based on a better approximation of the image formation model, it reaches a superior rectification quality. This is especially visible in Figure 7a, where both the road and railway are still distorted when correlation is used, but little to no deformation is visible with our approach.

Second, as our approach is based on the Bayesian formalism, it includes a prior for the latent variable. This prior in turn has a regularizing effect (Jospin et al., 2022), which prevents the solution from collapsing to a slightly more optimal shift according to the image texture, but less likely in practice. This effect is especially visible in Figure 7b, where the texture on the road across the image seems to have just the right tilt to cause strong and visible deformations on the other elements of the image, here a tree. With our method, no such distortion is apparent on the rectified image.

4.2 Tie point detection and filtering

Here, we investigate whether increasing the accuracy of the horizontal shift estimate increases the quality of detected tie points. We ran our tie point detector, first on the unrectified data, second on the data rectified using the correlation based approach described by (Fang et al., 2017) and third on the data rectified using our method. We also used the reference GPS/INS data to rectify the hyperspectral data and manually selected 949 tie points. These manual tie points will be used later as comparison points when evaluating the accuracy of the boresight calibration.

The tie points obtained by the three different methods are manually categorized as either inliers, when visibly pointing at the same object, or outliers, when pointing to different objects. Around 1'000 tie points were labeled in this way for each set of points. This allows for the different methods to be compared in accuracy in terms of the proportion of inliers present in the tie points.

Finally, we used the homography based RANSAC filtering approach described in Section 3.4 to filter the outliers. We estimate the proportion of inliers a second time, after filtering.

The results, reported in Table 1, show clearly that using a y-scaling invariant scheme, as we propose, significantly improves both rectification methods by increasing the number of tie points and the proportion of inliers. Additionally, the proposed filtering method based on a homography model yields a larger proportion of inliers, but is not perfect. In fact, horizontal shift without y-scaling invariance seems to be counterproductive compared to running the detector on the raw data, most

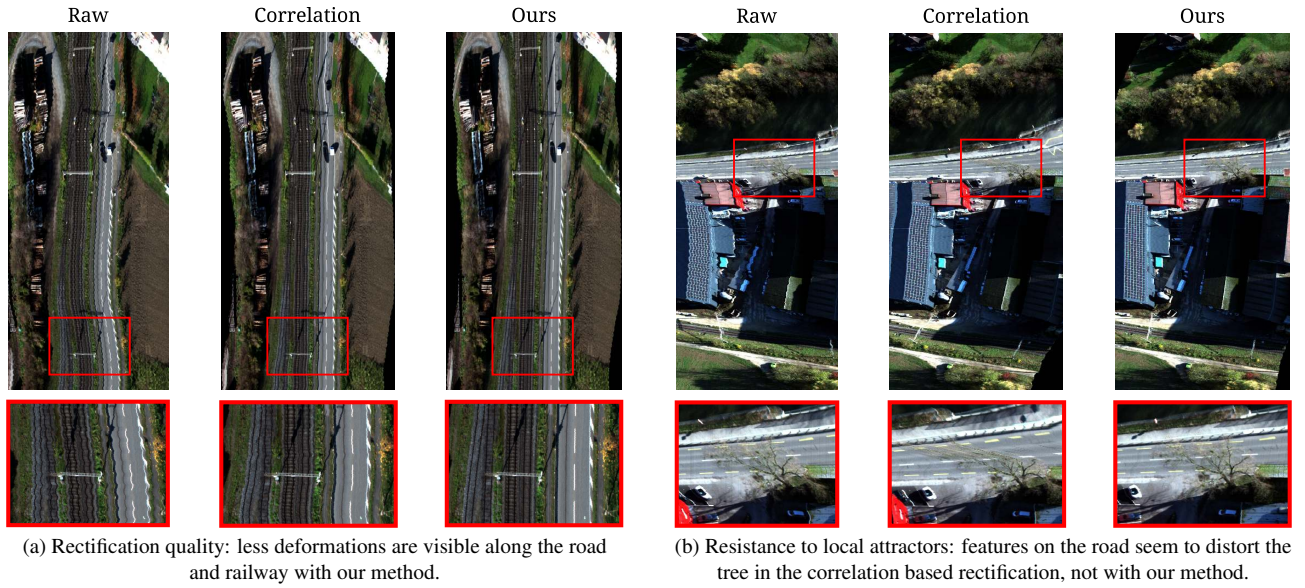


Figure 7. Our method is both more accurate than the state of the art (a), but also, the Bayesian prior ensures that the solution is not too strongly impacted by local attractors (b)

Method	Y-scale invariant	Keypoints	Initial matches (Inliers prop)	Selected Ransac (actual Inliers prop)
Reference	-	-	(manual) 949 (100.0%)	-
Raw	No	288'024	2'213 (16.0%)	871 (acc. 69.9%)
(Fang et al., 2017)	No	310'433	3'088 (12.5%)	1'145 (acc. 69.6%)
(Fang et al., 2017)	Yes	500'971	12'691 (58.8%)	7'920 (acc. 92.2%)
Ours	No	309'187	2'380 (15.0%)	917 (acc. 72.7%)
Ours	Yes	499'947	14'459 (69.2%)	10'289 (acc. 94.0%)

Table 1. Number of tie points and inliers for different methods. Bold is the best estimated proportion of inliers (excluding reference)

likely because the horizontal correction re-aligns patterns on the road, which in turn increases the amount of self-similar texture areas where the algorithm could fail. Our method reaches a higher proportion of inliers, both before (70%) and after filtering (95%), showing the benefits of an improved horizontal rectification algorithm.

Overall, our method produces a higher proportion of inlier tie points within the highest number of matches. In the next section, we investigate if this higher proportion of inlier tie points translates to a better estimation of the boresight calibration, which is an important practical application we are interested in.

4.3 Camera boresight calibration

To estimate the uncertainty associated with the obtained set of tie points, we use bootstrapping (Efron and Tibshirani, 1985), i.e., estimate the uncertainty associated with an input distribution by using resampling in the set of measurements. While Laplace's approximation (MacKay, 2003) is often used to determine the uncertainty of an estimate obtained via Gauss-Newton's optimization (the covariance of the estimated posterior is just the inverse of the Hessian of the optimization problem), it also assumes that the errors in the inputs are normally distributed. Laplace's approximation is also meant to estimate the uncertainty given a set of measurements, not the uncertainty associated with the generation of the measurements themselves. To account for the fact that the error distribution of the obtained tie points can be heavily tailed and assumed to be arbitrary, we use bootstrapping instead.

For each set of filtered points shown in Table 1, including the

949 reference points, we randomly select 500 and solve the problem of boresight calibration. For each tie point set, except for the reference which we know contains no outliers, we compare the results when using either the ℓ_2 or the Huber kernel for the Gauss-Newton optimization.

We repeat this experiment 100 times. From the 100 generated samples, we compute the mean, as well as the expected angle deviation from the mean. The angle deviation between two rotations $\mathbf{r}_1$ and $\mathbf{r}_2$, with $\mathbf{r}_1$ and $\mathbf{r}_2$ expressed as axis angles, can be computed as:

$$\arccos \left(\frac{\text{Tr} \left(e^{-[\mathbf{r}_1]_{\times}} e^{[\mathbf{r}_2]_{\times}} \right) - 1}{2} \right). \quad (8)$$

We then compare the boresight estimate for each generated tie point set and compare it with the estimate obtained with manual tie points (which we will use as ground truth), computing the deviation with Equation 8 again.

The results, shown in Table 2, highlight the clear benefits of our method. We get a 30% reduction in the uncertainty associated with the tie point generation process compared to the method of (Fang et al., 2017), and a 50% reduction in the error compared to the ground truth (i.e., the reference obtained with manual tie points). The caveat is that the use of a robust regression kernel like the Huber loss, seems to be mandatory, as without, the remaining outliers seem to have an out-sized effect, but with less than 10% outliers after RANSAC, the robust estimator remains stable.

Method	Filter method	Optimization kernel	Bootstrap standard error	Error with reference
Reference	Manual	ℓ_2	0.17°	-
Raw	Homography Ransac	ℓ_2	5.77°	10.35°
Raw	Homography Ransac	Huber	1.33°	5.72°
(Fang et al., 2017)	Homography Ransac	ℓ_2	4.92°	20.97°
(Fang et al., 2017)	Homography Ransac	Huber	0.30°	0.34°
Ours	Homography Ransac	ℓ_2	3.59°	11.37°
Ours	Homography Ransac	Huber	0.22°	0.12°

Table 2. Comparison of different boresight calibration approaches, with expected angular error and error compared to the reference. Bold is the best accuracy (excluding reference)

5. Conclusion

In this work, we have demonstrated that a method for push-broom hyperspectral horizontal shift rectification based on probabilistic modeling of the image formation process, coupled with a y-scale invariant tie point extraction algorithm and robust regression effectively addresses the boresight calibration problem and achieves accuracy on par with manually curated tie points. In addition to the calibration of the boresight for light-weight hyperspectral push-broom cameras, our method for tie point generation can be used for other problems such as integrated orientation of GPS/INS or other re-calibration procedures for the camera parameters.

Since our method for tie point production does not require external input, another interesting direction for future research is multiple sensor co-registration.

6. Acknowledgements

Julien Burkhard conducted most of the experiments described in this paper under the supervision of Jesse Lahaye and Laurent Jospin. Experiments were designed by Laurent Jospin. Processing tools for hand labelling of the ground truth data were developed by Laurent Jospin. The overall research was directed by Jan Skaloud. Thanks to Julien Vallet and Sixense Helimap for facilitating the acquisition of our dataset. Thanks to Davide Gucci for their valuable inputs.

This project was supported by the Open Research Data Program of the ETH Board, by ORD 5th contribute call, project OGAIS, which covered the development of the manual labeling tools used for obtaining the ground truth data used in our experiments, and ORD 8th contribute call, project OPAS, which covered the travel cost and part of the development of the automatic tie points system.

References

Abdel-Aziz, Y. I., Karara, H. M., Hauck, M., 2015. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Photogrammetric engineering & remote sensing*, 81(2), 103–107. [4](#)

Agisoft, 2024. Metashape. <https://www.agisoft.com/>. Accessed: 2024-07-31. [1](#)

Alcantarilla, P. F., Solutions, T., 2011. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell.*, 34(7), 1281–1298. [2, 4](#)

Asner, G. P., 1998. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment*, 64(3), 234–253. [1](#)

Barbieux, K., Constantin, D., Merminod, B., 2016. Correction of Airborne Pushbroom Images Orientation Using Bundle Adjustment of Frame Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B3, 813–818. [1, 2](#)

Berveglieri, A., Tommaselli, A. M. G., 2019. Geometric Filtering of Matches Between Points in Bands of Hyperspectral Cubes. *IEEE Geoscience and Remote Sensing Letters*, 16(3), 492–496. [2](#)

Buntine, W. L., 1994. Operations for Learning with Graphical Models. *Journal of Artificial Intelligence Research*, 2, 159–225. [3](#)

Ding, A., Liang, S., Jiao, Z., Ma, H., Kokhanovsky, A. A., Peltoniemi, J., 2022. Improving the Asymptotic Radiative Transfer Model to Better Characterize the Pure Snow Hyperspectral Bidirectional Reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16. [1](#)

Dozier, J., Painter, T. H., 2004. Multispectral and Hyperspectral Remote Sensing of Alpine Snow Properties. *Annual Review of Earth and Planetary Sciences*, 32(Volume 32, 2004), 465–494. [1](#)

Efron, B., Tibshirani, R., 1985. The Bootstrap Method for Assessing Statistical Accuracy. *Behaviormetrika*, 12(17), 1–35. [6](#)

Fang, H., Hu, B., Hao, S., Wu, Y., Liu, Y., Li, A., 2017. Two-step rectification method for distortions caused by aircraft roll in push-broom hyperspectral images. *Journal of Applied Remote Sensing*, 11(4), 046019–046019. [1, 2, 5, 6, 7](#)

Genton, M. G., 2002. Classes of kernels for machine learning: a statistics perspective. *J. Mach. Learn. Res.*, 2, 299–312. [3](#)

Habib, A., Zhou, T., Masjedi, A., Zhang, Z., Evan Platt, J., Crawford, M., 2018. Bore-sight Calibration of GNSS/INS-Assisted Push-Broom Hyperspectral Scanners on UAV Platforms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1734–1749. [2](#)

Hartley, R., 1997. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6), 580–593. [1](#)

He, H., Siu, W.-C., 2011. Single image super-resolution using gaussian process regression. *CVPR 2011*, 449–456. [3](#)

Hueni, A., Geier, S., Vögtli, M., LaHaye, J., Rosset, J., Berger, D., Sierro, L. J., Jospin, L. V., Thompson, D. R., Schläpfer, D., Green, R. O., Loeliger, T., Skaloud, J., Schaepman, M. E., 2025. The AVIRIS-4 Airborne Imaging Spectrometer. *IEEE Geoscience and Remote Sensing Letters*, 1–1. [2](#)

Jensen, R. R., Jackson, M. W., Lulla, V., 2008. Single line correction method to remove aircraft roll errors in hyperspectral imagery. *Journal of Applied Remote Sensing*, 2(1), 023529. [2](#)

Jospin, L. V., Laga, H., Boussaid, F., Buntine, W., Bennamoun, M., 2022. Hands-On Bayesian Neural Networks—A Tutorial for Deep Learning Users. *IEEE Computational Intelligence Magazine*, 17(2), 29–48. [5](#)

Lacey, A., Pinitkarn, N., Thacker, N. A., 2000. An evaluation of the performance of ransac algorithms for stereo camera calibration. *BMVC*, Citeseer, 1–10. [4](#)

Liu, D. C., Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. *Mathematical programming*, 45(1), 503–528. [3](#)

Liu, L., Coops, N. C., Aven, N. W., Pang, Y., 2017. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sensing of Environment*, 200, 170–182. [1](#)

MacKay, D. J., 2003. *Information theory, inference and learning algorithms*. Cambridge university press. [6](#)

Nex, F., Armenakis, C., Cramer, M., Cucci, D., Gerke, M., Honkavaara, E., Kukko, A., Persello, C., Skaloud, J., 2022. UAV in the advent of the twenties: Where we stand and what is next. *ISPRS Journal of Photogrammetry and Remote Sensing*, 184, 215–242. [1](#), [2](#)

Ordonez, A., Arguello, F., B. Heras, D., 2018. Alignment of Hyperspectral Images Using KAZE Features. *Remote Sensing*, 10, 756. [4](#)

Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. Orb: An efficient alternative to sift or surf. *2011 International Conference on Computer Vision*, 2564–2571. [4](#)

Ryan R. Jensen, Andrew J. Hardin, P. J. H., Jensen, J. R., 2011. A New Method to Correct Pushbroom Hyperspectral Data Using Linear Features and Ground Control Points. *GIScience & Remote Sensing*, 48(3), 416–431. [2](#)

SA, P., 2024. Professional photogrammetry and drone mapping software — pix4d. <https://www.pix4d.com/>. Accessed: 2024-07-31. [1](#)

Sankararao, A. U., Kumar, N. S., Rajalakshmi, P., 2020. Workflow and calibration of airborne hyperspectral imaging system. *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 757–762. [1](#), [2](#)

Sharma, A., Gilgien, S., Skaloud, J., 2024. Model-based fusion of gnss and multiple-imus. *2024 11th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, 266–271. [1](#)

Tareen, S. A. K., Saleem, Z., 2018. A comparative analysis of sift, surf, kaze, akaze, orb, and brisk. *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 1–10. [4](#)

Turner, D., Lucieer, A., McCabe, M., Parkes, S., Clarke, I., 2017. Pushbroom Hyperspectral Imaging from an Unmanned Aircraft System (UAS) – Geometric Processing Workflow and Accuracy Assessment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W6, 379–384. [2](#)

Yu, Y., Ma, Y., Mei, X., Fan, F., Huang, J., Ma, J., 2021. A Spatial-Spectral Feature Descriptor for Hyperspectral Image Matching. *Remote Sensing*, 13(23), 4912. Number: 23 Publisher: Multidisciplinary Digital Publishing Institute. [2](#)

Zhang, Z., 2021. Epipolar constraint. *Computer Vision: A Reference Guide*, Springer, 387–387. [1](#)