Published in Image Processing On Line on 2015–12–26.
Submitted on 2015–07–11, accepted on 2015–12–07.
ISSN 2105–1232 © 2015 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials,
software, datasets and online demo at
https://doi.org/10.5201/ipol.2015.146

Attitude Refinement for Orbiting Pushbroom Cameras: a Simple Polynomial Fitting Method

Carlo de Franchis[1], Enric Meinhardt-Llopis[1], Daniel Greslou[2], Gabriele Facciolo[1]

[1] CMLA, ENS Cachan, CNRS, Université Paris-Saclay, France
({carlo.de-franchis, enric.meinhardt, facciolo}@cmla.ens-cachan.fr)
[2] CNES, Toulouse, France (daniel.greslou@cnes.fr)

Communicated by Pierre Moulon *Demo edited by* Carlo de Franchis

Abstract

This paper describes a simple pushbroom camera model for Earth observation satellites and proposes a new algorithm to refine the orientation parameters of a camera from a set of ground control points. The relative importance of the various orientation parameters is analyzed. On the last generation of high resolution satellites such as Pléiades and WorldView, the attitude angles are shown to be the main contributors to localization errors. Thus the proposed algorithm focuses on refining the attitude angles. It is based on a simple polynomial fitting method. Numerous experiments, which can be reproduced through the online demo associated to this paper, show that the proposed algorithm is able to reduce the localization error by one order of magnitude with only a few ground control points. A geometric simulator for the proposed model is implemented, as well as the attitude refinement algorithm.

Source Code

The reviewed Python source code implementing a geometric simulator for the camera model and the attitude refinement algorithm is available from the web page of this article[1]. Usage instructions and documentation are included in the `README.md` file of the archive.

Keywords: pushbroom sensor; satellite orientation; attitude refinement; ground control points

1 Introduction

The number of optical Earth observation satellites has increased drastically over the past decade. According to the World Meteorological Organization[2], 65 civilian observation satellites are currently in operation, while at the end of the Nineties they were around 30. Their proliferation will certainly

[1] https://doi.org/10.5201/ipol.2015.146

[2] Observing Systems Capability Analysis and Review Tool, http://www.wmo-sat.info/oscar/satellites (accessed on 8 july 2015)

accelerate as new companies such as Skybox plan to launch dozens of low-cost satellites in the next few years. Starting during the Fifties, the development of the first Earth observation satellites was mostly driven by military needs related to the Cold War, but civilian applications followed very quickly in fields such as meteorology, cartography, geology, hydrology, forestry or glaciology. Nowadays, images from Earth observation satellites are used to solve problems ranging from geographic mapping to meteorology predictions [18], including applications as diverse as the measurement of elevation changes for glaciers [2] or rescue assistance for natural disasters [43].

Most of these applications require the images to be accurately georeferenced. For instance when using multi-temporal images of the same area to monitor changes in vegetation, water bodies, or geological landmarks, the images have to be accurately registered in order to provide meaningful measurements. Obviously for this use case a relative registration of the images may suffice to the end users, but it implies a preprocessing specific to the application. In an ideal world, all satellite images would be natively georeferenced to sub-pixel precision.

To georeference an image, two things are necessary. First an accurate geometric model of the camera with all the numerical values of its parameters, and second a digital surface model of the ground area corresponding to the viewed scene. This paper deals with the first thing. Observation satellites use pushbroom cameras which record the scene line by line as the satellite moves forward on its orbit. The resulting image is a stack of lines taken at slightly different times, from different positions and viewing angles. The camera position and viewing direction in space (usually referred to as *attitude*) are recorded on board during the image acquisition. The frequency and precision of these measurements depend on the satellite, but even in the best cases the resulting accuracy is limited and needs to be refined using the image content. A common solution is to use known correspondences between *ground control points*, i.e. 3D points located in the viewed scene, and their projection on the image. The camera parameters are then estimated as those that best fit the correspondences according to the geometric camera model.

1.1 Previous Work

Many papers describing pushbroom camera models have been published since the early Fifties and the advent of aerial photography. A thorough analysis of these works was made recently by Poli and Toutin in their last review paper [31], and also by Konecny [19] and Joseph [18] in their respective textbooks. An exhaustive review was already made in 2004 by Toutin [38]. We give here a quick historical overview of the developments on pushbroom camera models and estimation algorithms, but the interested reader should refer to these publications for a comprehensive coverage.

Papers dealing with pushbroom cameras were published even before the launch of the first observation satellites, as pushbroom cameras were used on aircrafts [4]. In these papers pushbroom cameras are generally referred to as *strip cameras*.

The first camera model for orbiting pushbroom cameras was published in 1983 by Guichard and Toutin [10, 35]. The model was used for SPOT-1 which was the first civilian Earth observation satellite to be launched after the American Landsat series. Later this model became known as the *Toutin Model* and was embedded in the commercial software PCI Geomatica. This model assumes that the satellite flies on a perturbed elliptical orbit and that its nominal constant attitude is perturbed by second degree polynomial functions of time. Later the model was extended and adapted to newer observation satellites [36, 37, 39, 40].

In 1987, Konecny proposed to model the movement of orbiting pushbroom cameras with a straight path and a constant attitude, which leads to a model with six parameters [20]. The differences between that model and reality were handled by eight perturbation parameters. This model was extended later by Jacobsen [14, 15, 16].

In 1987, Gugan proposed an orbiting pushbroom camera model for SPOT-1 in which the satellite

flies an elliptical orbit whose longitude of the ascending node changes linearly with time, as well as the satellite true anomaly [9]. The attitudes are assumed to drift linearly with time, and their initial value is unknown. This model was extended by Michalis and Dowman in 2008 [24].

In 1988, Kratky proposed to model the attitude variations of SPOT-1 by quadratic functions of time [21], [23]. This model was implemented in the software SPOTCHECK+ and was later extended by Fritsch and Stallmann [8].

In 1990, Westin proposed to simplify the usual models by approximating the orbit of SPOT-1 by a plane circular orbit [42]. This approximation has to be valid only during the timespan of one scene. He reported experiments on SPOT-1 images showing that the ground error due to this approximation is less than 10 cm. To further simplify the model, the attitude derivatives measured on board are assumed to be accurate enough, leaving only a constant term (i.e. the initial attitude) to be estimated. An iterative least-squares minimization was proposed to estimate the seven parameters of the model (four for the circular orbit and three for the constant attitude term). Sub-pixel accuracies were obtained on SPOT-1 images using a limited number of control points.

In 1995, Gupta and Hartley described a model similar to Westin's model, in a paper aimed at describing a new algorithm for mapping 3D points to the image plane [11]. The only difference of their model with respect to Westin's is that the satellite path is assumed to be elliptic instead of circular. But Gupta and Hartley do not explain how the orbit parameters and the attitude offset are estimated from ground control points. In 1997 they introduced and gave a rigorous analysis of the *linear pushbroom camera* [12]. This model is considerably simpler than the usual models because it assumes that the satellite flies at a uniform speed on a straight line and that its attitude is constant. As a consequence, the relationship between image and ground coordinates is linear in terms of homogeneous coordinates, almost as with standard pinhole cameras. The authors reported a maximal error of half a pixel on a whole SPOT-1 scene when comparing the linear pushbroom model to the more realistic model they published two years earlier [11].

In 2005, Poli proposed to model both the position and attitude errors by second order piecewise polynomials depending on time [28, 29, 30]. For pushbroom cameras carried by satellites, two segments were reported to be sufficient in general [29]. The position and attitude errors are used to correct the cubic splines interpolated from the position and attitude samples measured on board and provided by the image vendors.

In 2007, Crespi et al. proposed to model the satellite path by an elliptical orbit and the attitude errors by second order polynomials depending on time [5]. The measured attitude functions are obtained by interpolation from samples provided by vendors. This model was implemented in the SISAR software, but the implementation is not distributed. The whole model was detailed and analyzed in 2009 in a book chapter [6].

1.2 Our Contribution

Poli and Toutin grouped the major Earth observation satellites in several classes and subclasses [31]. The present paper focuses on the most recent subclass of satellites, which scan the Earth in *asynchronous* mode and use a so-called *agile configuration*. These satellites fly along sun-synchronous and quasi-polar orbits and provide very high resolution images using a flexible single-lens optical system. They have the ability to rotate about the axes of their camera in order to point at off-nadir targets up to 30°, or to view the target from different directions during a single orbit. Some well-known commercial satellites of this category are Pléiades 1A and 1B, QuickBird, WorldView 1, 2, and 3, GeoEye 1, Eros A and B, Orbview 3 and TopSat.

This paper proposes a simple model for this kind of orbiting pushbroom cameras and analyzes the main causes for localization inaccuracies. A simple algorithm is proposed to improve the localization accuracy. To the best of our knowledge, this paper is the first published work that provides an open

source implementation of an orbiting pushbroom camera simulator as well as an online demo allowing to test the proposed attitude refinement algorithm.

The next section exposes the camera model used in the paper. Section 3 analyzes the main causes for localization inaccuracies and proposes an algorithm to reduce them. Section 4 presents numerous synthetic experiments to validate the proposed approach.

2 Pushbroom Cameras

This section describes the pushbroom camera model used in the paper. This model is an extension of the model that was described by Westin in 1990 [42]. The main difference is that our model is built for satellites whose attitude vary continuously during the timespan of one scene, while the models made for SPOT-1 [42, 10, 35, 20, 9, 21, 11] were based on the assumption that the attitude was unknown but almost constant. In our model the attitude evolution is thus described by polynomial functions of degree 3.

2.1 The Linear Pushbroom Camera

The goal of a camera model is to define the geometric relationship between image and space coordinates. We will first recall how to derive this relationship with the specialized case of the *linear pushbroom camera* that was introduced by Gupta and Hartley in 1997 [12]. The linear pushbroom camera model relies on two simplifying assumptions:

- the camera is traveling in a straight line at a constant velocity with respect to the world,
- the attitude of the camera is constant over the whole image acquisition duration.

This camera can be thought of as a standard pinhole camera [13] moving at constant speed along a straight line in 3-space, with fixed attitude. This pinhole camera would have the additional property of containing a single line of pixel sensors in its focal plane instead of the usual bi-dimensional array of pixel sensors. Thus, at any moment, only the points lying in the plane defined by the optical center and the line of pixel sensors are imaged by the camera. This plane is called the *view plane*, and is shown in Figure 1. The orientation of the view plane depends only on the camera attitude, and is assumed to be fixed. It is also assumed that the motion of the camera does not lie in the view plane. Consequently, the view plane sweeps out the whole 3-space as time varies. The image provided by such a camera is made of lines scanned independently at different instants of time and stored one next to the other.

All the coordinates in Figure 1 are expressed in the instantaneous camera frame, defined below.

Definition 1. *The* camera frame *is an orthonormal coordinate frame attached to the moving camera. The origin of the coordinate frame is the instantaneous camera center. The* Z*-axis lies in the instantaneous view plane and is perpendicular to the linear sensor array. It is directed towards the visible half-space. The* X*-axis is orthogonal to the view plane and its direction is determined by requiring that the motion of the camera has a positive* X *component. The* Y*-axis lies in the view-plane, is parallel to the sensor array, and its direction is determined by requiring that the* X, Y*, and* Z *axes form a right-handed coordinate frame.*

The image of a 3-space point $\mathbf{X}$ is described by its row and column coordinates x and y. The first coordinate x gives the time instant at which the point $\mathbf{X}$ was imaged, i.e. lay in the view plane. If δ_t denotes the time spent by the sensor to measure one sample, then the instant of time at which

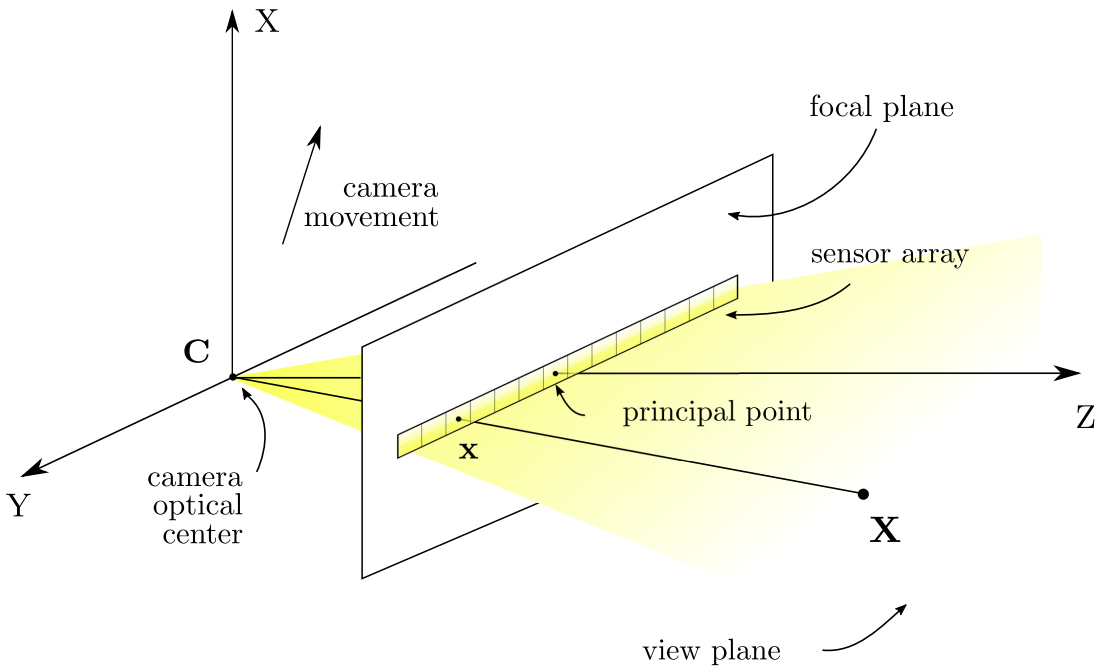


Figure 1: Linear pushbroom camera model. The instantaneous camera optical center **C** moves at a constant speed. The view plane sweeps out the whole 3-space as the camera moves forward. The 3D point **X** is imaged by the camera to the focal plane location **x**.

the row x was imaged is $x\delta_t$ (assuming that the first row was imaged at $t = 0$). The duration δ_t is called the *dwell time*. It is part of the internal parameters of the sensor.

The second coordinate y gives the position of the projection of **X** on the sensor. It is a 2D projection carried out in the view plane. The coordinates of the projected point can be expressed in the camera frame. In that frame the view plane is defined by $\mathrm{X} = 0$. If f denotes the *focal length* of the camera, it follows from similar triangles that the point $\mathbf{X} = (0, \mathrm{Y}, \mathrm{Z})^\top$ is mapped to the point $(0, f\mathrm{Y}/\mathrm{Z}, f)^\top$, as shown in Figure 2.

The coordinate y is a pixel index expressed in the sensor frame, which differs from the camera frame by a translation. The magnitude of this translation is given by the *principal point* coordinate y_0. It is the pixel index of the projection of the camera center on the sensor array, as shown on Figure 3. If w denotes the *pixel size* measured with the camera frame units, then $w(y - y_0) = f\mathrm{Y}/\mathrm{Z}$.

This permits to describe the locus of 3-space points that may have been imaged on a given image point.

Result 1. *Let* $\mathbf{x} = (x, y)^\top$ *be an image point, where* x *is the row index and* y *the column index. Take as coordinate frame the instantaneous camera frame at time* $t = x\delta_t$*. Then the inverse image of* $\{\mathbf{x}\}$ *by the linear pushbroom camera mapping is the 3-space line passing through the origin and directed by* $(0, w(y - y_0), f)^\top$*.*

2.2 Camera Rotation

Most modern high resolution commercial satellites such as those of the Pléiades and WorldView series scan the Earth in *asynchronous mode* [31]. The word *asynchronous* refers to the fact that the satellite speed and ground scanning speed are not equal: the satellite moves faster than the projection of its sensor array on the ground. On the one hand this ensures a longer exposure time, which allows more photons to reach the CCD (Charge Coupled Device) sensor and thus improves the signal to noise

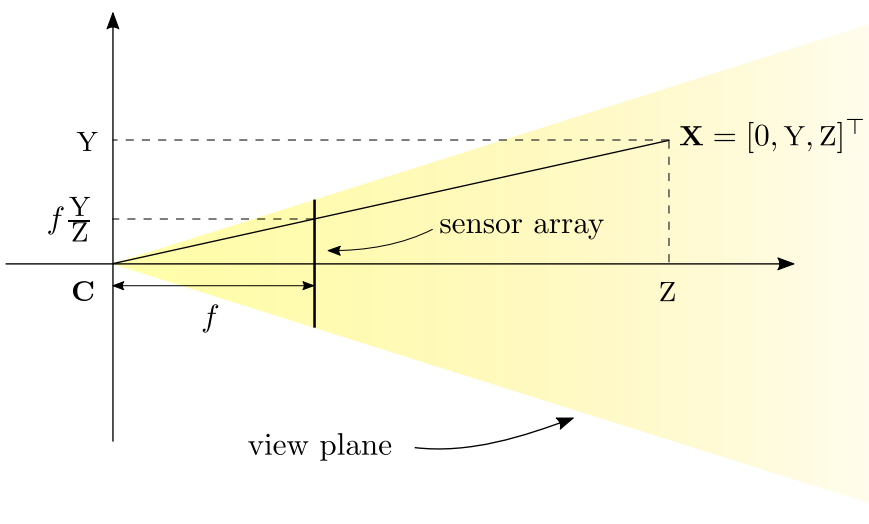


Figure 2: For 3-space points in the view plane, the image formation works in the same way as with a pinhole camera. From similar triangles one knows that the ratio between the Y coordinate of the image point and f is equal to $\frac{Y}{Z}$.

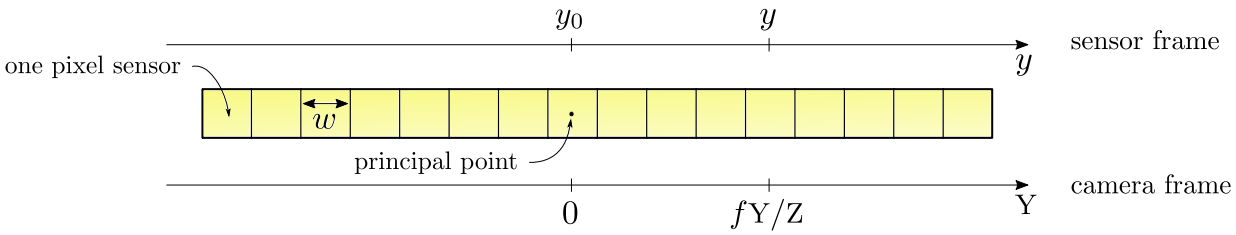


Figure 3: The sensor frame is used to express the coordinate of a pixel in a given image row. Its relationship with the camera frame depends on internal parameters such as the pixel size w and the principal point coordinate y_0.

ratio [28]. On the other hand, it implies that the camera attitude is not constant during one image acquisition, i.e. that the second assumption made by the linear pushbroom camera model is not valid, as illustrated in Figure 4. This makes the image formation model slightly more complicated and increases the number of parameters needed to describe the image acquisition system.

The first consequence of the variation of attitude angles over time on our camera model is that the camera frame has no longer a fixed orientation. To convert coordinates between the camera frame and a fixed world frame, an intermediate coordinate frame is thus needed. This frame is the *local orbital frame*.

Definition 2. *The* local orbital frame *is an orthonormal coordinate frame attached to the moving satellite. The origin of the coordinate frame is the satellite center of mass. The* Z*-axis points towards the center of the Earth. The* X*-axis is defined by the motion of the satellite. The* Y*-axis is determined by requiring that the* X, Y*, and* Z *axes form an orthonormal right-handed coordinate frame.*

For the sake of simplicity the camera center is supposed to coincide with the satellite center of mass. The camera frame and the local orbital frame have thus the same origin, and differ only by a rotation. This rotation is described by its roll, pitch and yaw components (φ, ψ, ω). To obtain the camera frame, the local orbital frame is first rotated about its X axis by the roll angle φ, then rotated about its new axis Y′ by the pitch angle ψ, and finally rotated about the twice new axis Z″ by the yaw angle ω. These three rotations are shown on Figure 5. As the three elemental rotations

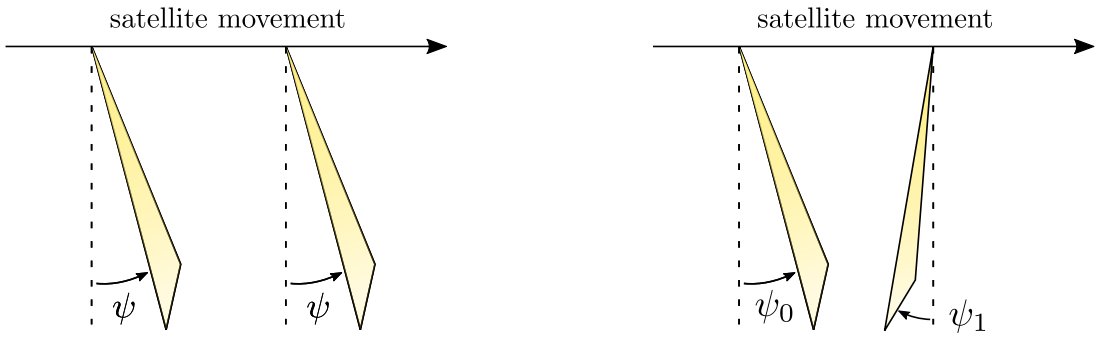


Figure 4: *Synchronous* observation satellites, shown on the left, have fixed attitudes: the orientation of the camera with respect to the satellite does not change over time. On *asynchronous* observation satellites, shown on the right, the attitude angles may vary during the image acquisition, such as the pitch angle ψ shown here.

occur about axes of the rotating coordinate system (they are called *intrinsic rotations*), the resulting transformation matrix is obtained by composing the associated *Givens* rotation matrices [13] from left to right

$$\mathtt{R}(\varphi,\psi,\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\varphi & -s_\varphi \\ 0 & s_\varphi & c_\varphi \end{bmatrix} \begin{bmatrix} c_\psi & 0 & s_\psi \\ 0 & 1 & 0 \\ -s_\psi & 0 & c_\psi \end{bmatrix} \begin{bmatrix} c_\omega & -s_\omega & 0 \\ s_\omega & c_\omega & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where c_a and s_a denote $\cos(a)$ and $\sin(a)$. Note that the values of the attitude angles φ, ψ, ω depend on t. The change of basis matrix from the local orbital frame to the camera frame will usually be written as $\mathtt{R}(\varphi(t), \psi(t), \omega(t))$.

Result 2. *Let $\mathbf{x} = (x, y)^\top$ be an image point, where x is the row index and y the column index. In the local orbital frame at time $t = x\delta_t$, the inverse image of $\{\mathbf{x}\}$ by the pushbroom camera mapping is the 3-space line passing through the origin and directed by*

$$\mathtt{R}(\varphi(t), \psi(t), \omega(t)) \begin{pmatrix} 0 \\ w(y - y_0) \\ f \end{pmatrix}. \tag{2}$$

Our model makes the extra assumption that the roll, pitch and yaw are polynomial functions of time of degree 3

$$\begin{aligned} \varphi(t) &= \varphi_0 + \varphi_1 t + \varphi_2 t^2 + \varphi_3 t^3 \\ \psi(t) &= \psi_0 + \psi_1 t + \psi_2 t^2 + \psi_3 t^3 \\ \omega(t) &= \omega_0 + \omega_1 t + \omega_2 t^2 + \omega_3 t^3. \end{aligned}$$

The attitude of the camera is thus entirely described by the set of coefficients of these 3 polynomial functions, which is denoted by $\Phi = \{(\varphi_k, \psi_k, \omega_k)\}_{k=0,\dots,3}$. Additional terms may also be considered to model certain oscillations of the satellite attitude [26], but this is beyond the scope of the present paper.

2.3 Satellite Movement

Real satellites have elliptical orbits around the Earth as described by Kepler's laws. Hartley and Gupta [12] reported a discrepancy of about 5 m when approximating an orbiting pushbroom with a linear pushbroom camera model on images from the first generation of SPOT satellites.

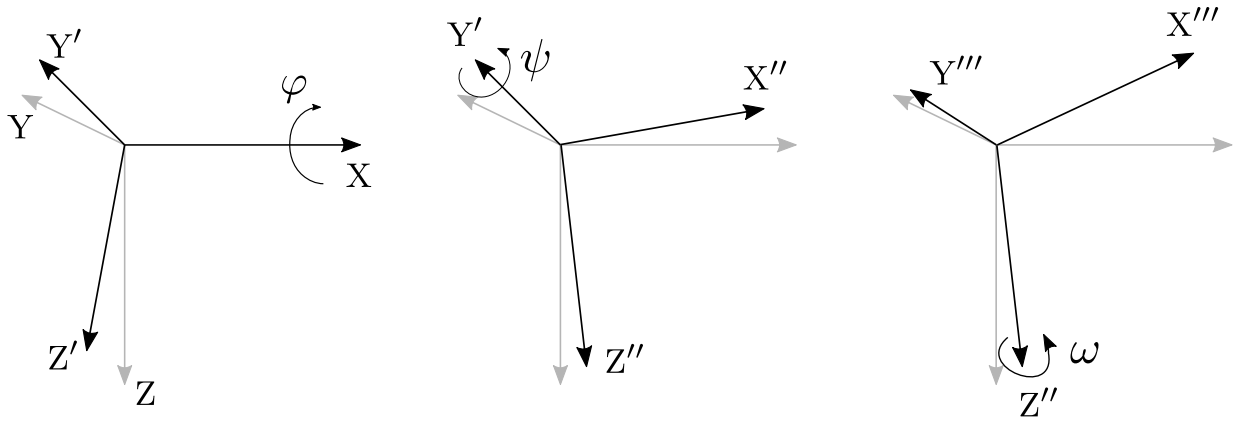


Figure 5: The camera frame (represented in black) is obtained from the local orbital frame (in gray) by composing three *Givens* rotations [13] about the three axes of the rotating frame. The first rotation is about the X axis, thus the X axis does not change and $\mathrm{X}' = \mathrm{X}$. In the same way $\mathrm{Y}'' = \mathrm{Y}'$ and $\mathrm{Z}''' = \mathrm{Z}''$. The roll angle φ shown on the left figure is negative, while the pitch and yaw angles ψ and ω shown on the middle and right figures are positive.

As the first SPOT satellites had constant attitudes, the main factor responsible for this discrepancy might be the first assumption of the linear pushbroom camera model, i.e. the uniform and linear camera movement hypothesis.

Orbit description. In this paper the orbit of the satellite is assumed to be a circle centered on the Earth's center of mass, as in [42]. It is described by three parameters: the radius of the circle, plus two angles giving the orientation of the plane in which the circle lies.

The radius is $R_E + a$, where R_E denotes the Earth radius and a the flying altitude of the satellite, usually around 700 km. The orientation of the orbital plane is given by its *inclination* $i \in [0, \pi]$ and its *reference node longitude* $\lambda_0 \in [-\pi, \pi]$. The inclination is the angle formed with the equatorial plane, as shown on Figure 6. The reference node is the intersection between the orbit and the equatorial plane at which the satellite is ascending (i.e. going towards north). The longitude λ_0 is measured with respect to the *inertial frame*.

Definition 3. *The* Earth-centered inertial frame *is an orthonormal coordinate frame attached to the center of mass of the Earth, whose axes have fixed directions. Its* Z *axis is the Earth rotation axis, oriented towards the North pole. Its* X *axis lies in the equatorial plane and is permanently fixed in a direction relative to the celestial sphere. The* Y*-axis is determined by requiring that the* X, Y*, and* Z *axes form an orthonormal right-handed coordinate frame.*

The inertial frame does not rotate with the Earth. This means that any particular location on Earth, except the two poles, has coordinates that change over time. Coordinates of points on Earth are usually expressed in a frame that rotates with the Earth, namely the *Earth centered rotational frame*, defined below.

Definition 4. *The* Earth centered rotational frame*, or* rotational frame *for short, is an orthonormal coordinate frame attached to the rotating Earth. The origin of the coordinate frame is the center of mass of the Earth. The* Z*-axis is pointing towards the north. The* X*-axis intersects the surface of the Earth at* 0° *latitude (equator) and* 0° *longitude (Greenwich). The* Y*-axis is determined by requiring that the* X, Y*, and* Z *axes form an orthonormal right-handed coordinate frame. The coordinates of a point in that frame are called its* geographic coordinates*. They are usually given in spherical form by the* longitude λ*, the* latitude θ *and the* altitude h *above the surface of the sphere.*

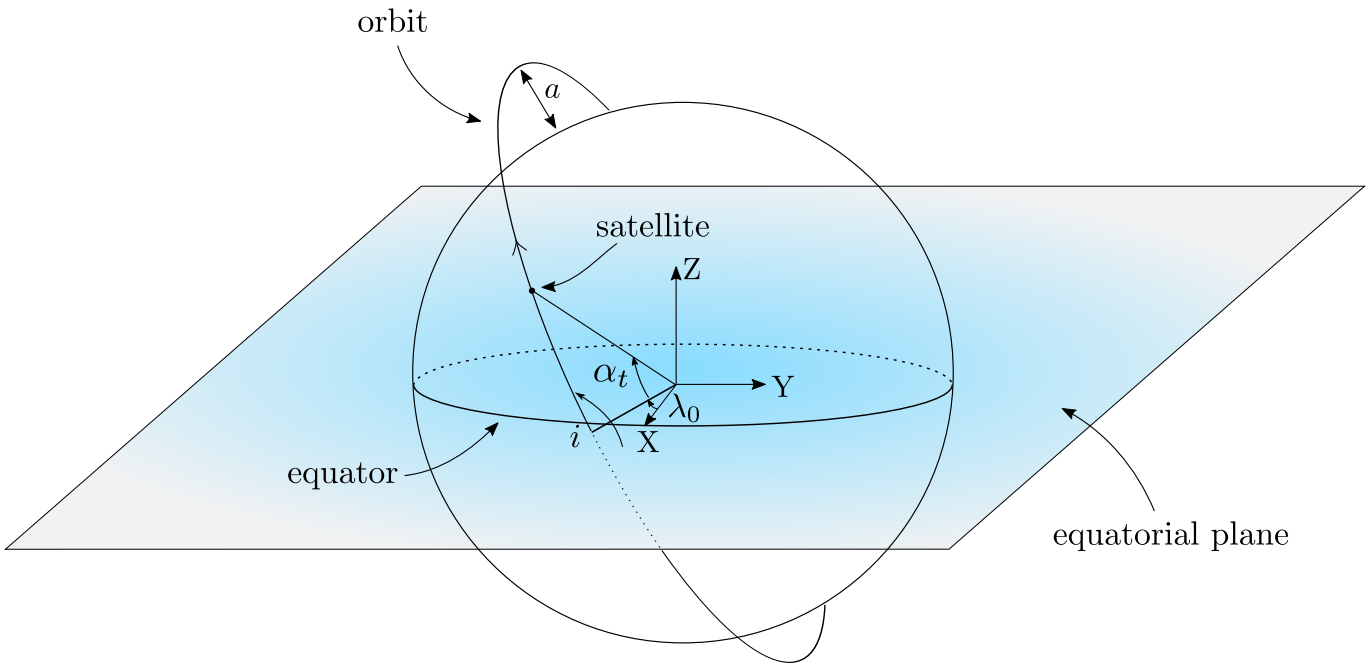


Figure 6: A circular orbit around the Earth is described by three parameters: its radius $R_E + a$, the longitude λ_0 of its reference node, and its inclination i. The longitude λ_0 is defined with respect to the **X** axis of the inertial frame. The inclination i is the angle between the equatorial plane and the orbital plane. The satellite is localized by its angular position α_t. Note that the longitude represented here is negative.

For the sake of simplicity, we assume that the rotational frame coincides with the inertial frame at the beginning of the acquisition ($t = 0$). The frame rotates with the Earth, therefore coordinates of a fixed point on the Earth surface do not change over time.

Position of the satellite. The position of the satellite on the orbit is given by an angle $\alpha_t \in [0, 2\pi]$. As the orbit is supposed to be circular and uniform, α_t depends only on its initial value α_0 and on the time t elapsed since the beginning of the acquisition

$$\alpha_t = \alpha_0 + 2\pi \frac{t}{T_S}, \tag{3}$$

where T_S denotes the orbital period of the satellite. It is known from Kepler's third law that

$$\frac{2\pi}{T_S} = \sqrt{\frac{\mu}{(R_E + a)^3}}, \tag{4}$$

where μ is the standard gravitational parameter of the Earth, thus T_S depends only on a. For typical Earth observation satellites with a roughly equal to 700 km, T_S is about one hour and a half.

Change of basis. The change of basis matrix from the rotational frame to the local orbital frame depends both on the orbit parameters and the satellite position. As the rotational frame and the satellite position depend on the time t, this change of basis matrix depends on t. It is a simple rotation composed with a translation. Its expression is derived in the next two paragraphs.

From inertial to orbital frame. Given a satellite position α_t and a triplet of orbit parameters (λ_0, i, a), the change of basis is obtained by first rotating the inertial frame about its Z axis by λ_0.

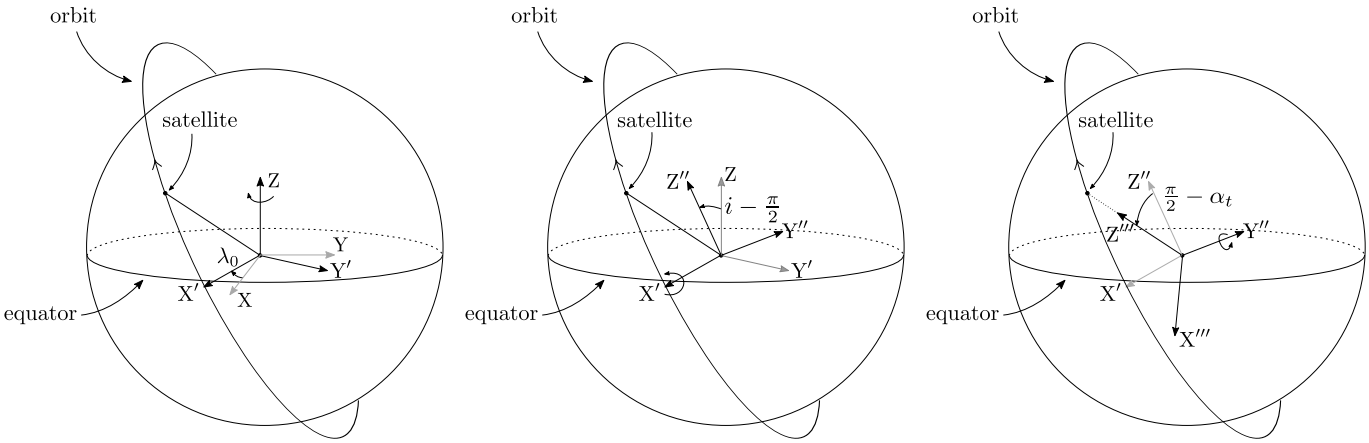


Figure 7: Decomposition of the rotation that maps the axes of the Earth-centered inertial frame to the axes of the local orbital frame. The rotation is obtained by composing three *Givens* rotations about the three axes of the rotating frame. In each figure, the old axes are plotted in gray while the new axes are plotted in black. In the first figure, the inertial frame is rotated about its Z axis by λ_0. In the second figure, the new axes are rotated about X′ by $i - \frac{\pi}{2}$. In the last figure, the new axes are rotated about Y″ by $\frac{\pi}{2} - \alpha_t$.

This makes the new X axis, X′, point to the reference node, as shown in Figure 7. The new frame is then rotated about X′ to send the Z axis in the orbital plane. As the Z axis is perpendicular to the equatorial plane, the angle between the X′Z plane and the orbital plane is $i - \frac{\pi}{2}$. A third rotation by $\frac{\pi}{2} - \alpha_t$ about the new Y axis, Y″, makes the Z axis point towards the satellite. An additional rotation of π about the same axis makes the Z axis point towards the Earth center. The change of basis matrix is obtained by composing these three rotations

$$\mathtt{P}_{\text{inertial, orbital}}(\lambda_0, i, \alpha_t) = \begin{bmatrix} c_{\lambda_0} & -s_{\lambda_0} & 0 \\ s_{\lambda_0} & c_{\lambda_0} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{i-\frac{\pi}{2}} & -s_{i-\frac{\pi}{2}} \\ 0 & s_{i-\frac{\pi}{2}} & c_{i-\frac{\pi}{2}} \end{bmatrix} \begin{bmatrix} c_{-\alpha_t-\frac{\pi}{2}} & 0 & s_{-\alpha_t-\frac{\pi}{2}} \\ 0 & 1 & 0 \\ -s_{-\alpha_t-\frac{\pi}{2}} & 0 & c_{-\alpha_t-\frac{\pi}{2}} \end{bmatrix} \tag{5}$$

From rotational to inertial frame. The change of basis at time t is a rotation of angle $\tau = 2\pi\frac{t}{T_E}$, where T_E denotes the duration of a stellar day, which is the Earth rotation period relative to the fixed stars. It is approximately 86 164.10 s. The axis of the rotation is Z, thus the change of basis matrix is

$$\mathtt{P}_{\text{rotational, inertial}}(t) = \begin{bmatrix} c_\tau & -s_\tau & 0 \\ s_\tau & c_\tau & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

The change of basis matrix from rotational to orbital frame is obtained by composing the previous matrices

$$\begin{aligned} \mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0) &= \mathtt{P}_{\text{rotational, inertial}}(t)\mathtt{P}_{\text{inertial, orbital}}(\lambda_0, i, \alpha_0 + 2\pi\frac{t}{T_S}) \\ &= \begin{bmatrix} c_{\tau+\lambda_0} & -s_{\tau+\lambda_0} & 0 \\ s_{\tau+\lambda_0} & c_{\tau+\lambda_0} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{i-\frac{\pi}{2}} & -s_{i-\frac{\pi}{2}} \\ 0 & s_{i-\frac{\pi}{2}} & c_{i-\frac{\pi}{2}} \end{bmatrix} \begin{bmatrix} c_{-\alpha_t-\frac{\pi}{2}} & 0 & s_{-\alpha_t-\frac{\pi}{2}} \\ 0 & 1 & 0 \\ -s_{-\alpha_t-\frac{\pi}{2}} & 0 & c_{-\alpha_t-\frac{\pi}{2}} \end{bmatrix}, \end{aligned} \tag{7}$$

where for readability α_t is used to denote $\alpha_0 + 2\pi\frac{t}{T_S}$ in the last matrix (see Equation 3).

Result 3. *Let $\mathbf{x} = (x, y)^\top$ be an image point. In the Earth centered rotational frame at time $t = x\delta_t$, the inverse image of $\{\mathbf{x}\}$ by the pushbroom camera mapping is the 3-space line passing through the point*

$$\mathbf{S} = -\mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0) \begin{pmatrix} 0 \\ 0 \\ R_E + a \end{pmatrix}, \tag{8}$$

and directed by the vector

$$\mathbf{u} = \mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0)\mathtt{R}(\varphi(t), \psi(t), \omega(t)) \begin{pmatrix} 0 \\ w(y - y_0) \\ f \end{pmatrix}. \tag{9}$$

Proof. Let us denote by $\mathbf{S}$ the center of mass of the satellite. From Result 2 we know that the inverse image of $\{\mathbf{x}\}$ is the 3-space line passing through $\mathbf{S}$ and directed by a vector whose coordinates in the orbital frame are

$$\mathtt{R}(\varphi(t), \psi(t), \omega(t))(0, w(y - y_0), f)^\top. \tag{10}$$

The coordinates of that vector in the rotational frame are obtained by composing with the change of basis matrix $\mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0)$, which leads to the expression of $\mathbf{u}$ given in (9). If we denote by $\mathbf{E}$ the Earth's center of mass, then $\mathbf{S} - \mathbf{E}$ is a vector whose coordinates in the orbital frame are $-(0, 0, R_E + a)^\top$. Thus its coordinates in the rotational frame are given by

$$-\mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0)(0, 0, R_E + a)^\top. \tag{11}$$

As $\mathbf{E}$ is the origin of the rotational frame, the coordinates of $\mathbf{S}$ in that frame are the same. □

2.4 Image Formation Model

Localization function. Given an image point $\mathbf{x}$ and the altitude h above the sea level of the associated 3-space point $\mathbf{X}$, it is now possible to compute the geographic coordinates of $\mathbf{X}$ (see Definition 4). This is done by computing the first intersection point of the line defined in Result 3 with the sphere of radius $R_E + h$ centered at the Earth center.

The line defined in Result 3 passes through the point $\mathbf{S}$ and is directed by the vector $\mathbf{u}$ given by expression (9). The first intersection point with the sphere is computed by finding a scalar r such that

$$\|\mathbf{S} + r\mathbf{u}\|^2 = (R_E + h)^2. \tag{12}$$

This is a polynomial equation on r of degree 2. When the camera points towards the Earth, this equation has two positive roots, and the smallest root

$$r = \frac{-\mathbf{S}^\top\mathbf{u} - \sqrt{(\mathbf{S}^\top\mathbf{u})^2 - \mathbf{u}^\top\mathbf{u}(\mathbf{S}^\top\mathbf{S} - (R_E + h)^2)}}{\mathbf{u}^\top\mathbf{u}} \tag{13}$$

gives the position of the desired 3-space point $\mathbf{X} = \mathbf{S} + r\mathbf{u}$. This defines the *localization* function

$$\begin{aligned} L : \mathbf{R}^2 \times \mathbf{R} &\to [-\pi, \pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \\ (\mathbf{x}, h) &\mapsto (\lambda, \theta), \end{aligned} \tag{14}$$

which associates a longitude λ and a latitude θ to each image point $\mathbf{x}$, provided that the altitude h is known. Here λ and θ are the geographic coordinates of the point $\mathbf{X} = \mathbf{S} + r\mathbf{u}$.

Projection function. Conversely, given a 3-space point $\mathbf{X}$ it is possible to compute the coordinates of its image $\mathbf{x}$. As there is no direct way of computing the time t at which the view plane will sweep through $\mathbf{X}$, this computation is not as straightforward as the previous one. This is a notable difference with the pinhole camera model for which the computation of image coordinates is a linear operation on homogeneous coordinates of the 3-space point. The *projection function*

$$\begin{aligned} P : [-\pi, \pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times \mathbf{R} &\to \mathbf{R}^2 \\ (\lambda, \theta, h) &\mapsto \mathbf{x}, \end{aligned} \tag{15}$$

is the inverse of L with respect to the image coordinates. It is estimated iteratively from L. An initial estimate is computed by assuming that the dependence with h is linear and the dependence with (λ, θ) is quadratic. This initial estimate is then refined using the Newton algorithm.

A first estimation P_0 of the projection function P is made for a fixed altitude h. The estimation is done by assuming that the function $P(\cdot, \cdot, h)$ is quadratic. A 2×6 matrix $\mathtt{P}$ is estimated such that for all λ, θ

$$P_0(\lambda, \theta, h) = \mathtt{P}(1, \lambda, \theta, \lambda^2, \theta^2, \lambda\theta)^\top. \tag{16}$$

The matrix $\mathtt{P}$ is obtained through least squares minimization from a list of correspondences

$$(x_i, y_i) \leftrightarrow (\lambda_i, \theta_i) = L(x_i, y_i, h), \tag{17}$$

obtained by evaluating the localization function L on the vertices of a regular sampling grid $(x_i, y_i)_{i=1,\dots,n^2}$ of size $n \times n$ in the image domain (usually $n = 5$ is enough). The two rows $\mathbf{P}^{1\top}$ and $\mathbf{P}^{2\top}$ of $\mathtt{P}$ are computed by least-squares minimization of respectively $\|\mathtt{A}\mathbf{p} - \mathbf{b}_1\|$ and $\|\mathtt{A}\mathbf{p} - \mathbf{b}_2\|$, where

$$\mathtt{A} = \begin{bmatrix} 1 & \lambda_1 & \theta_1 & \lambda_1^2 & \theta_1^2 & \lambda_1\theta_1 \\ & & & \vdots & & \\ 1 & \lambda_{n^2} & \theta_{n^2} & \lambda_{n^2}^2 & \theta_{n^2}^2 & \lambda_{n^2}\theta_{n^2} \end{bmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_{n^2} \end{pmatrix}, \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix} y_1 \\ \vdots \\ y_{n^2} \end{pmatrix}. \tag{18}$$

The previous least-squares estimation is carried out for two different altitudes h_1 and h_2, leading to two matrices $\mathtt{P}_1$ and $\mathtt{P}_2$. These two altitudes can be, for example, lower and upper bounds for the 3D points of the scene. The projection function estimate P_0 is finally obtained from linear interpolation on h

$$P_0(\lambda, \theta, h) = \left(\frac{|h_2 - h|}{|h_2 - h_1|}\mathtt{P}_1 + \frac{|h - h_1|}{|h_2 - h_1|}\mathtt{P}_2 \right) (1, \lambda, \theta, \lambda^2, \theta^2, \lambda\theta)^\top. \tag{19}$$

Note that the estimation of $\mathtt{P}_1$ and $\mathtt{P}_2$ is done only once, while Equation (19) has to be evaluated each time we want to evaluate the projection function. The initial estimate P_0 is then refined using the Newton method [32], as detailed in Algorithm 1. At each iteration, the function $(x, y) \mapsto L(x, y, h)$ is approximated by its first order Taylor expansion. The inversion of its Jacobian matrix, denoted by $J_{x,y}L(x, y, h)$, allows to refine the estimation. Note that the partial derivatives needed to compute the Jacobian matrix are approximated by forward differences. For example:

$$\frac{\partial L}{\partial x}(x, y, h) \approx L(x + 1, y, h) - L(x, y, h). \tag{20}$$

The loop stops when the function $L(\cdot, \cdot, h)$ maps the refined image coordinates to a point on Earth which is less than 1 mm from the input. This distance is computed using the fact that the squared distance between two points with spherical coordinates (λ, θ) and $(\tilde{\lambda}, \tilde{\theta})$ on the unit sphere is approximated to first order by $(\tilde{\theta} - \theta)^2 + (\tilde{\lambda} - \lambda)^2 \cos^2\theta$ when $(\tilde{\lambda}, \tilde{\theta})$ is close to (λ, θ).

Algorithm 1: Iterative estimation of the projection function $P : (\lambda, \theta, h) \mapsto (x, y)$.

Input: λ, θ, h: geographic coordinates and altitude of a 3D point
n_{iter}: number of iterations
Output: x, y: image coordinates
begin

$x, y = P_0(\lambda, \theta, h)$ *// initialization*
for $k \in \{1 \dots n_{\text{iter}}\}$ **do**
  $\tilde{\lambda}, \tilde{\theta} = L(x, y, h)$
  if $R_E \sqrt{(\tilde{\theta} - \theta)^2 + (\tilde{\lambda} - \lambda)^2 \cos^2 \theta} < 1\,\text{mm}$ *// error check*
  then
    break
  $\begin{pmatrix} x \\ y \end{pmatrix} \leftarrow \begin{pmatrix} x \\ y \end{pmatrix} - [J_{x,y} L(x, y, h)]^{-1} \begin{pmatrix} \tilde{\lambda} - \lambda \\ \tilde{\theta} - \theta \end{pmatrix}$ *// update*

	Parameter	Description	Pléiades	WorldView-2
intrinsics	δ_t	dwell time	0.07 ms	?
	w	pixel width	13 µm	8 µm
	f	focal length	12.9 m	13.3 m
	y_0	principal point coordinate	15000 px	17500 px
extrinsics	a	orbit altitude	694 km	770 km
	i	orbit inclination	98.2°	98.5°
	λ_0	reference node longitude	−180° - 180°	−180° - 180°
	α_0	initial angular position on the orbit	0° - 360°	0° - 360°
	Δ_t	duration of the acquisition	≈ 3 s	≈ 3 s
	$\Phi = (\varphi_k, \psi_k, \omega_k)_{k=0,\dots,3}$	coefficients of the attitude functions		

Table 1: Parameters of the pushbroom camera model. The last two columns give typical numerical values for Pléiades and WorldView-2 satellites. The parameters may be divided into two classes: *internal parameters*, sometimes called *intrinsics*, which characterize the camera, and *external parameters*, or *extrinsics*, which characterize the position and orientation of the camera with respect to the inertial frame (see Definition 3).

The projection and localization functions depend obviously on the values of the camera parameters. For the sake of simplicity, this dependence is not made explicit in Equation (14) but this may be done if needed by renaming L to $L_{\delta_t, w, f, y_0, a, i, \lambda_0, \alpha_0, \Phi}$. The complete set of 20 parameters on which L depends (12 of them are grouped together into Φ), plus the duration of the acquisition Δ_t, is summarized in Table 1.

3 Pushbroom Camera Parameters Refinement

The computation of the camera parameters, both external and internal, from a set of world to image correspondences is a problem known as *resectioning* [13]. In the case of remote sensing images, the internal parameters are usually known with a good accuracy because the sensor has been carefully calibrated before the launch of the satellite. Moreover the sensor calibration may have been refined during the in-flight commissioning, as happened for Pléiades [7]. In this work the internal parameters are thus assumed to be perfectly known.

3.1 Camera Position Parameters

On modern high resolution satellites, the external parameters are measured on board in real time and provided with the images. The attitude parameters are more sensitive to measurement inaccuracies than the position parameters: for the typical case of a nadir looking satellite orbiting at 700 km, an error of 1 µrad on the roll or pitch angles yields a displacement of about 70 cm on the ground, while an error of 1 cm on the camera position yields a displacement of at most 1 cm on the ground. The accuracy of the attitude is much more critical than that of the camera position. On Pléiades the orbit parameters, i.e. the satellite positions, are computed on board with centimeter accuracy by two DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite) positioning instruments [7]. Note that satellite trajectories are subject to much less perturbations than airplanes trajectories.

Moreover, in terms of ground coordinates, attitude and position errors have similar effects and are therefore difficult to distinguish from each other. A roll or pitch error has locally the same effect as a translation of the focal plane, which can be obtained through a translation of the camera center. This observation was already reported by Gugan [9], Konecny [20] and Kratky [23] in the late Eighties about SPOT-1 orientation parameters.

Therefore we make the extra assumption that the camera positions (i.e. the path followed in 3-space by the optical center of the camera) are perfectly known, and that all the measurement inaccuracies lie in the attitude angles.

3.2 Camera Attitude Parameters

The problem tackled in this section is to estimate the values of the 12 attitude parameters $\Phi = (\varphi_k, \psi_k, \omega_k)_{k=0,\dots,3}$, given a set of correspondences between image and world points. To emphasize the fact that only the attitude parameters are being estimated, the localization and projection functions L and P introduced in Section 2.4 are denoted by L_Φ and P_Φ.

3.2.1 Yaw versus Roll and Pitch

If we assume that the satellite is at nadir, i.e. with zero roll and pitch angles, the effect of attitude errors on the localization function L_Φ can be computed simply, as shown on Figure 8. To first order, a small error ε on the roll or pitch angles induces a displacement of $a\varepsilon$ on the ground, where a is the distance between the satellite and the ground. Things are different for the yaw, as it induces a rotation of the imaging sensor around its center. A small error ε results in a displacement of at most $\frac{D}{2}\varepsilon$ on the ground, where D is the swath width of the satellite. The relative importance of yaw errors with respect to roll and pitch errors is thus $\frac{D}{2a}$, which is generally less than 2%. For example with Pléiades $D = 20$ km and $a = 694$ km, the ratio is approximately 1.4%.

It is thus reasonable to neglect the yaw error and assume that the yaw coefficients are perfectly known. The problem is thus narrowed down to estimating the value of the roll and pitch coefficients given a set of world to image correspondences. For simplicity of notation, the reduced set of 8 attitude parameters $(\varphi_k, \psi_k)_{k=0,\dots,3}$ is still denoted by Φ.

3.2.2 Global Energy Minimization

Assume $n \geq 1$ point correspondences $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ between 3D points $\mathbf{X}_i$ and 2D image points $\mathbf{x}_i = (x_i, y_i)$ are given. x_i and y_i are the row and column indices of the image points, meaning that $x_i\delta_t$ is the acquisition time and y_i is the column index, i.e. the position on the pushbroom array. When the attitude coefficients Φ are exact, the line $L_\Phi(\mathbf{x}_i, \cdot)$ passes through the 3-space point $\mathbf{X}_i$ and the energy

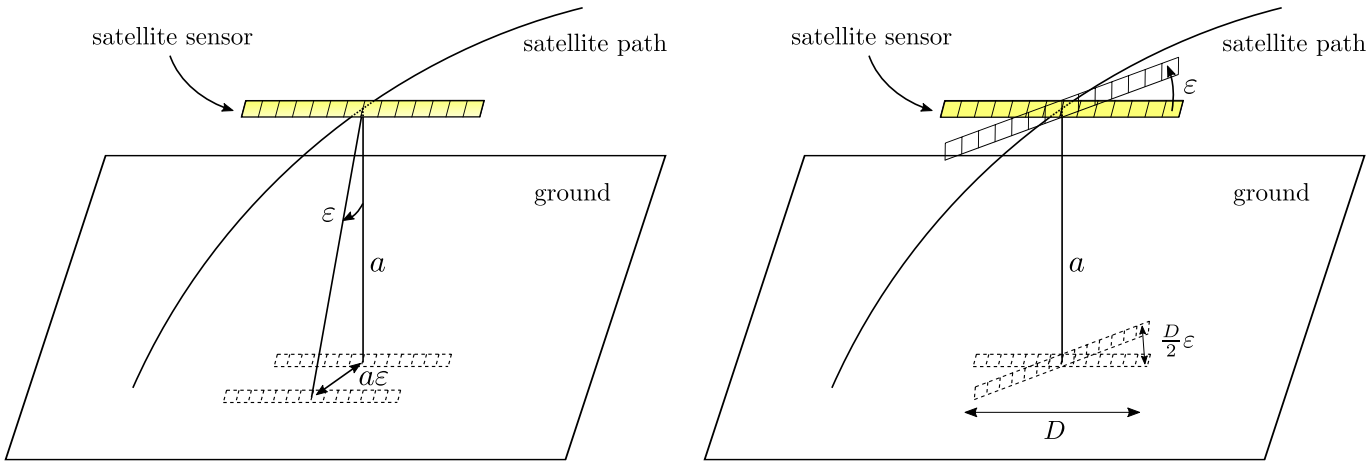


Figure 8: Effect of attitude errors on the localization function. The figure on the left shows a pitch error of ε. To first order, this induces a ground displacement of the sensor projection of $a\varepsilon$, where a is the flying altitude of the sensor. The figure on the right shows a yaw error of ε. It induces a maximal ground displacement of $\frac{D}{2}\varepsilon$, obtained for the sensor endpoints, where D denotes the swath width.

E defined by

$$E(\Phi) = \frac{1}{n}\sum_{i=1}^{n} d(\mathbf{X}_i, L_\Phi(\mathbf{x}_i, \cdot))^2, \tag{21}$$

is zero. Here $d(\mathbf{X}, \ell)$ denotes the Euclidean distance between the 3-space point $\mathbf{X}$ and the line ℓ. Note that even if the line $L_\Phi(\mathbf{x}_i, \cdot)$ can be thought of as the set of points $L_\Phi(\mathbf{x}_i, h)$ obtained by evaluating L_Φ on $(\mathbf{x}_i, h)$ for h varying in $\mathbf{R}$, the line is entirely defined by a point and a direction (given in Result 3). Hence the evaluation of $L_\Phi(\mathbf{x}_i, h)$ for any particular h is not needed. The attitude parameters can then be estimated by

$$\hat{\Phi} = \underset{B(\Phi,\eta)}{\arg\min}\, E, \tag{22}$$

where Φ is the initial estimate measured on board, $B(\Phi, \eta)$ is the ball centered in Φ with radius η, and η is given by the nominal precision of the satellite attitudes measurement system (typically 50 µrad, i.e. around 30 m on the ground).

The cost expressed in Equation (21) involves the distance between 3-space points $\mathbf{X}_i$ and rays back-projected from image points $\mathbf{x}_i$. Another option could be to minimize the distance, in the image plane, between image points $\mathbf{x}_i$ and projected points $P_\Phi(\mathbf{X}_i)$:

$$E'(\Phi) = \frac{1}{n}\sum_{i=1}^{n} d(\mathbf{x}_i, P_\Phi(\mathbf{X}_i))^2, \tag{23}$$

This would require to evaluate the projection function P_Φ instead of the localization function L_Φ, which, as reported by Gugan [9] and Kratky [22], is more complex computationally since each evaluation of P_Φ requires several evaluations of L_Φ and its derivatives (see Algorithm 1). For solving the general camera resectioning problem, energy (23) is usually preferred to energy (21) because in energy (21) the points located far away from the camera have more importance than the points located close to the camera. But in the case of remote sensing images all the distances between points are small compared to the distance separating the camera from the scene. The two energies can thus be considered as equivalent.

3.2.3 Pointwise Minimization

An alternative approach to global minimization is to compute, for any correspondence $\mathbf{X} \leftrightarrow \mathbf{x}$, the instantaneous roll and pitch angles $\varphi(t)$ and $\psi(t)$ needed to make the back-projected ray from $\mathbf{x}$, namely $L(\mathbf{x}, \cdot)$, pass through $\mathbf{X}$. The direction of this particular back-projected ray depends only on the values of the roll and pitch angles at time $t = x\delta_t$, not on the whole set of coefficients Φ. The coefficients in Φ are then estimated by polynomial fitting from the n computed roll and pitch values.

Each world to image correspondence imposes instantaneous roll and pitch values. Let us consider a single correspondence between the image point $\mathbf{x} = (x, y)$ and the world point $\mathbf{X}$. Knowing the instantaneous yaw $\omega(t)$ at time $t = x\delta_t$, we look for roll and pitch values φ and ψ such that the ray back-projected from $\mathbf{x}$ passes through $\mathbf{X}$. Assume that the coordinates of $\mathbf{X}$ are expressed in the local orbital frame at time t (the conversion from geographic coordinates to local orbital coordinates is done thanks to Equation (7)). Using the direction of the ray back-projected from $\mathbf{x}$ given in Result 2, our problem amounts to find φ and ψ such that $\mathtt{R}(\varphi, \psi, \omega(t))(0, w(y - y_0), f)^\top$ has the same direction as $\mathbf{X}$. The factorization of $\mathtt{R}(\varphi, \psi, \omega)$ given in Equation (1) allows to separate the known yaw term from the unknown roll and pitch: $\mathtt{R}(\varphi, \psi, \omega) = \mathtt{R}(\varphi, \psi, 0)\mathtt{R}(0, 0, \omega)$. By denoting

$$\mathbf{u} = \mathtt{R}(0, 0, \omega(t)) \begin{pmatrix} 0 \\ w(y - y_0) \\ f \end{pmatrix} = \begin{bmatrix} c_{\omega(t)} & -s_{\omega(t)} & 0 \\ s_{\omega(t)} & c_{\omega(t)} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ w(y - y_0) \\ f \end{pmatrix}, \tag{24}$$

the problem becomes finding φ and ψ such that

$$\mathtt{R}(\varphi, \psi, 0)\mathbf{u} \propto \mathbf{X}. \tag{25}$$

With normalized coordinates for $\mathbf{u}$ and $\mathbf{X}$ the proportionality constraint is replaced by an equality. The angles φ and ψ are obviously bounded to (at least) $[-\frac{\pi}{2}, \frac{\pi}{2}]$, otherwise the camera would not be oriented toward the Earth. Another reasonable assumption is that the Z coordinate of $\mathbf{u}$ and $\mathbf{X}$ is much bigger than both their X and Y coordinates. Even the most agile satellites such as Pléiades and WorldView-3 do not usually point more than 30° away from nadir. With that assumption our problem is then completely solved by the next result.

Result 4. *Let* $\mathbf{u} = (u_1, u_2, u_3)^\top$ *and* $\mathbf{v} = (v_1, v_2, v_3)^\top$ *be two points on the unit sphere* S^2*, such that* $u_3 > |u_1| + |v_1|\sqrt{2}$ *and* $v_3 > |v_2| + |u_2|\sqrt{2}$*. Then there exists a unique pair of angles* $(\varphi, \psi) \in [-\frac{\pi}{4}, \frac{\pi}{4}]^2$ *such that*

$$\mathtt{R}(\varphi, \psi, 0)\mathbf{u} = \mathbf{v}, \tag{26}$$

where $\mathtt{R}(\varphi, \psi, 0)$ *is the rotation obtained by composing two* Givens *rotations about the* X *and* Y *axis as defined in Equation* (1)*.*

Proof. Using the factorization of Equation (1), Equation (26) becomes

$$\begin{bmatrix} c_\psi & 0 & s_\psi \\ 0 & 1 & 0 \\ -s_\psi & 0 & c_\psi \end{bmatrix} \mathbf{u} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\varphi & s_\varphi \\ 0 & -s_\varphi & c_\varphi \end{bmatrix} \mathbf{v}, \tag{27}$$

which is equivalent to the following system of equations

$$\begin{aligned} u_1 \cos\psi + u_3 \sin\psi &= v_1, \\ v_2 \cos\varphi + v_3 \sin\varphi &= u_2, \\ -u_1 \sin\psi + u_3 \cos\psi &= -v_2 \sin\varphi + v_3 \cos\varphi. \end{aligned} \tag{28}$$

Lemma 1, stated below, shows that each of the first two independent equations has a unique solution on $[-\frac{\pi}{4}, \frac{\pi}{4}]$. Let us show that these solutions ψ and φ satisfy the third equation. As $\mathbf{u}$ and $\mathbf{v}$ lie on the unit sphere, one can find (α, β) and (γ, δ) in $[-\pi, \pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ such that

$$\begin{aligned} u_1 &= \cos\beta\cos\alpha & v_2 &= \cos\delta\cos\gamma, \\ u_3 &= \cos\beta\sin\alpha & v_3 &= \cos\delta\sin\gamma, \\ u_2 &= \sin\beta & v_1 &= \sin\delta. \end{aligned} \tag{29}$$

α and β (resp. γ and δ) are spherical coordinates for $\mathbf{u}$ (resp. $\mathbf{v}$). As $u_3 > |u_1|$ and $v_3 > |v_2|$, one can require additionally that α and γ lie in $]\frac{\pi}{4}, \frac{3\pi}{4}[$. With these notations the system (28) is rewritten as

$$\begin{aligned} \cos\beta\cos(\alpha - \psi) &= \sin\delta, \\ \cos\delta\cos(\gamma - \varphi) &= \sin\beta, \\ \cos\beta\sin(\alpha - \psi) &= \cos\delta\sin(\gamma - \varphi). \end{aligned} \tag{30}$$

The square of the left term of the third equation is

$$\cos^2\beta\sin^2(\alpha-\psi) = \cos^2\beta(1-\cos^2(\alpha-\psi)) = \cos^2\beta - \cos^2\beta\cos^2(\alpha-\psi) = \cos^2\beta - \sin^2\delta, \tag{31}$$

where the last equality comes from the first equation of (30). Similarly, using the second equation, the square of the right term is

$$\cos^2\delta\sin^2(\gamma-\varphi) = \cos^2\delta(1-\cos^2(\gamma-\varphi)) = \cos^2\delta - \cos^2\delta\cos^2(\gamma-\varphi) = \cos^2\delta - \sin^2\beta, \tag{32}$$

hence the two terms have the same squares. As β is in $[-\frac{\pi}{2}, \frac{\pi}{2}]$, $\cos\beta$ is non-negative. Moreover, as $\psi \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ and $\alpha \in]\frac{\pi}{4}, \frac{3\pi}{4}[$, it is clear that $0 < \alpha - \psi < \pi$ hence $\sin(\alpha - \psi) > 0$. Similarly, $\cos\delta$ and $\sin(\gamma - \varphi)$ are non-negative thus the equality between squares suffices to prove the desired result. □

Lemma 1. *Let a, b, c be three real numbers such that $|a| + |c|\sqrt{2} < b$. Then the equation*

$$a\cos x + b\sin x + c = 0 \tag{33}$$

has a unique solution on $[-\frac{\pi}{4}, \frac{\pi}{4}]$.

Proof. Let $f : [-\frac{\pi}{4}, \frac{\pi}{4}] \to \mathbf{R}$ be the mapping $x \mapsto a\cos x + b\sin x + c$. This function is continuous and has values of opposite sign at the endpoints of its domain interval, since

$$f\left(-\frac{\pi}{4}\right) f\left(\frac{\pi}{4}\right) = \left(\frac{a-b}{\sqrt{2}} + c\right)\left(\frac{a+b}{\sqrt{2}} + c\right) = \left(\frac{a}{\sqrt{2}} + c\right)^2 - \frac{b^2}{2} < 0. \tag{34}$$

Thus from the intermediate value theorem [33] f has a root in $[-\frac{\pi}{4}, \frac{\pi}{4}]$. To prove the uniqueness, the case $a = 0$ being trivial, one can assume that $a > 0$ (or otherwise use $-f$). The derivative of f is $x \mapsto -a\sin x + b\cos x$. It has a unique zero and changes sign from positive to negative in $\arctan\frac{b}{a}$, which is outside the domain $[-\frac{\pi}{4}, \frac{\pi}{4}]$ since $\frac{b}{a} > 1$. Thus f is strictly increasing and the root is unique. □

In practice the solution to Equation (33), and hence to equations (28) and (26), is computed thanks to the change of variable $t = \sin x$. If x is a solution then t satisfies the quadratic equation

$$(a^2 + b^2)t^2 + 2bct + c^2 - a^2 = 0, \tag{35}$$

which has two distinct solutions since its discriminant has the sign of $a^2 + b^2 - c^2$ and we assumed that $c^2 < b^2$. The unique solution to Equation (33) is found by testing these two candidates. The procedure to compute φ and ψ from a world to image correspondence is summarized in Algorithm 2.

Algorithm 2: Computation of roll and pitch values from a world to image correspondence.

Input: $\mathbf{X}$: Cartesian coordinates of a control point in the Earth-centered rotational frame
x, y: image coordinates
δ_t, w, f, y_0: camera internal parameters
a, i, λ_0: orbit parameters
α_0: satellite initial position on the orbit
ω: yaw function

Output: $\varphi, \psi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$: roll and pitch values

begin

$\mathbf{v} \leftarrow \mathtt{P}_{\text{rotational, orbital}}(x\delta_t, \lambda_0, i, \alpha_0)^{-1}\mathbf{X} + (0, 0, R_E + a)$ *// conversion to orbital coordinates*

$\mathbf{v} \leftarrow \frac{\mathbf{v}}{\|\mathbf{v}\|}$

$\mathbf{u} \leftarrow \mathtt{R}(0, 0, \omega(x\delta_t))(0, w(y - y_0), f)^\top$ *// yaw application*

$\mathbf{u} \leftarrow \frac{\mathbf{u}}{\|\mathbf{u}\|}$

if $u_3 < |u_1|+|v_1|\sqrt{2}$ or $v_3 < |v_2|+|u_2|\sqrt{2}$ **then**

  break *// the ground control point is unusable*

else

  solve $\mathtt{R}(\varphi, \psi, 0)\mathbf{u} = \mathbf{v}$ for φ and ψ *// using result 4*

return φ, ψ

Polynomial fitting of roll and pitch error evolution. Once a list of roll and pitch samples has been computed from the set of world to image correspondences thanks to Algorithm 2, the polynomial coefficients of the roll and pitch functions are estimated by least-squares minimization.

More precisely, assume that a list of roll samples $\hat{\varphi}^1, \ldots, \hat{\varphi}^n$ has been computed at times $t_1, \ldots, t_n$ from the world to image correspondences. If we denote by φ the initial roll function (the one derived from the onboard attitude measurements), then the roll correction function is estimated by fitting a polynomial of degree 3 to the points $t_i, \hat{\varphi}^i - \varphi(t_i)$ in $\mathbf{R}^2$. As the initial roll function has a known accuracy η, given by the satellite specifications, any true roll value cannot differ from $\varphi(t)$ more than η. The samples $t_i, \hat{\varphi}^i$ such that $|\hat{\varphi}^i - \varphi(t_i)| > \eta$ are thus discarded. Moreover the fitted correction polynomial must be bounded by η. Its coefficients $\mathbf{p} = (p_0, p_1, p_2, p_3)^\top$ are estimated by solving a constrained least squares problem

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \|\mathtt{A}\mathbf{p} - \mathbf{b}\|^2 \\ \text{subject to} \quad & -\eta \leq \sum_i p_i t^i \leq \eta, \, \forall t \in [0, T]. \end{aligned} \tag{36}$$

with

$$\mathtt{A} = \begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & t_n^3 \end{bmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} \hat{\varphi}^1 - \varphi(t_1) \\ \vdots \\ \hat{\varphi}^n - \varphi(t_n) \end{pmatrix}. \tag{37}$$

The constraint $-\eta \leq \sum_i p_i t^i \leq \eta, \, \forall t \in [0, T]$ is approximated thanks to a finite sampling $0, \frac{T}{m}, \frac{2T}{m}, \ldots, \frac{mT}{m}$ of the interval $[0, T]$, by

$$-\eta \leq \sum_i p_i \left(\frac{kT}{m}\right)^i \leq \eta \quad \forall k \in \{0, \ldots, m\}. \tag{38}$$

This linear constraint is expressed as

$$-\mathbf{h} \preceq \mathtt{G}\mathbf{p} \preceq \mathbf{h}, \tag{39}$$

where $\mathbf{x} \preceq \mathbf{y}$ denotes componentwise inequality, i.e. $x_i \leq y_i$ for all i, while $\mathtt{G}$ is an $(m+1) \times 4$ matrix and $\mathbf{h}$ an $(m+1)$-vector defined by

$$\mathtt{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & \frac{T}{m} & \frac{T}{m}^2 & \frac{T}{m}^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & T & T^2 & T^3 \end{bmatrix} \text{ and } \mathbf{h} = \begin{pmatrix} \eta \\ \vdots \\ \eta \end{pmatrix}. \tag{40}$$

The optimization problem (36) is thus replaced by a quadratic program problem [3]

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \|\mathtt{A}\mathbf{p} - \mathbf{b}\|^2 \\ \text{subject to} \quad & -\mathbf{h} \preceq \mathtt{G}\mathbf{p} \preceq \mathbf{h}. \end{aligned} \tag{41}$$

This kind of convex optimization problem that includes inequality constraints can be solved by interior-point methods. These methods solve the problem by applying Newton's method to a sequence of equality constrained problems [3]. A quadratic program solver is available in the `cvxopt` package, which is a free software package for convex optimization based on the Python programming language. It is developed by Martin Andersen, Joachim Dahl and Lieven Vandenberghe. This solver implements a primal-dual path-following method based on the Nesterov-Todd scaling [41, 1].

Note that for simplicity we detailed only the roll function estimation, but the method applies to pitch as well. The complete procedure to refine the roll and pitch functions is summarized in Algorithm 3.

Algorithm 3: Refinement of roll and pitch coefficients from world to image correspondences.

Input: $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$: set of $n \geq 1$ world to image correspondences
δ_t, w, f, y_0: camera internal parameters
a, i, λ_0: orbit parameters
α_0: satellite initial position on the orbit
φ, ψ, ω: attitude functions measured on board
η: nominal attitude accuracy
Output: $\hat{\varphi}, \hat{\psi}$: roll and pitch refined functions
begin

```
for k ∈ {1...n} do
    t_i ← x_i δ_t
    φ̂^i, ψ̂^i ← roll_and_pitch_from_gcp(X_i, x_i, δ_t, w, f, y_0, a, i, λ_0, α_0, ω)   // algorithm 2
    if |φ̂^i − φ(t_i)| > η or |ψ̂^i − ψ(t_i)| > η then
        discard φ̂^i, ψ̂^i, t_i, x_i, X_i
        n ← n − 1
A ← Vandermonde(t_1, ... t_n)                                   // Equation (37)
for θ ∈ {φ, ψ} do
    b ← (θ̂^i − θ(t_i))^⊤                                        // Equation (37)
    minimize_p ||Ap − b||^2 subject to −h ⪯ Gp ⪯ h               // G and h are precomputed
    θ̂ ← θ + Σ_i p_i t^i
return φ̂, ψ̂
```

4 Empirical Validation

This section presents an empirical validation procedure for the attitude correction Algorithm 3. All the experiments are carried out on synthetic data, using the pushbroom camera model described in Section 2. The implementation of that model is a *geometric*, and not *radiometric*, simulator.

On the one hand, synthetic data permits to test any particular configuration, enabling a fine understanding of the effects of each parameter. Moreover the synthetic observations, i.e. the 3D and 2D coordinates of world to image correspondences, are generated from the image formation model using the camera parameters, hence are noiseless. Adding artificial noise permits to analyze and quantify the impact of noise on the refinement algorithm.

On the other hand, the camera model used in the present paper is simplified: to mention just a few approximations, the orbit is supposed perfectly circular while real observation satellites fly on elliptical sun-synchronous orbits whose orbital plane precesses once a year, the Earth is supposed to be a sphere while a closer approximation would be an oblate spheroid, the light travel time from the Earth surface to the sensor is neglected, as well as the atmospheric refraction. Experiments on real data are obviously needed for a definitive validation of the proposed algorithm. This delicate task is left to a forthcoming article.

4.1 Experimental Setup

We use the same notational convention as Hartley and Zisserman in chapter 5 of their reference book [13]:

- measured quantities are denoted by $\mathbf{x}$ (image point coordinates), $\mathbf{X}$ (ground point coordinates) or Φ (attitude coefficients),
- estimated quantities are represented by a hat, such as $\hat{\mathbf{x}}$, $\hat{\mathbf{X}}$ or $\hat{\Phi}$,
- true values are represented by a bar, such as $\bar{\mathbf{x}}$, $\bar{\mathbf{X}}$ or $\bar{\Phi}$.

The testing procedure begins by choosing a fixed set of values for the parameters describing the pushbroom camera (see Table 1), among which the true attitude coefficients $\bar{\Phi}$ lie. Next a set of image locations $\bar{\mathbf{x}}_i$ and associated altitudes $\bar{h}_i$ are chosen. The number of such locations will vary. Synthetic world to image correspondences $\bar{\mathbf{X}}_i \leftrightarrow \bar{\mathbf{x}}_i$ are generated thanks to the localization function $L_{\bar{\Phi}}$. The correspondence is exact, in the sense that $\bar{\mathbf{X}}_i = L_{\bar{\Phi}}(\bar{\mathbf{x}}_i, \bar{h}_i)$ precisely, up to machine accuracy.

Next, artificial Gaussian noise is added to the image and ground locations by perturbing both the x and y coordinates of the image points and the X, Y and Z coordinates of the ground control points by a zero-mean Gaussian random variable with known variance. The resulting noisy points are denoted by $\mathbf{X}_i$ and $\mathbf{x}_i$. A suitable Gaussian random number generator is given in [32]. As the number of points is usually rather small, the experiment should be run many times with different realizations of the noise, and the results averaged. In the case of only one experiment, it makes more sense to impose the magnitude of the noise by choosing the perturbations uniformly on a sphere of known radius. This is what is implemented in Algorithm 4.

Next, the roll and pitch coefficients are perturbed. On real satellites, the attitude functions are interpolated from a list of regularly sampled attitude values. These attitude values are measured with star trackers and gyroscopes at fixed intervals. For examples, on WorldView-2, a state vector containing the time, position, velocity and attitude is recorded every 0.02 s [31]. On Pléiades, a state vector containing the time, position and velocity is recorded every 30 s during a timespan longer than the acquisition in order to provide around 10 samples. The attitudes are given directly as polynomial functions, but we can reasonably assume that they have been fitted from measured samples. To

simulate this behavior, $d+1$ real numbers are generated from a zero-mean uniform random variable with known amplitude. A polynomial perturbation of degree at most d is interpolated from these $d+1$ values, regularly spaced at times $t = 0, \frac{\Delta_t}{d}, \frac{2\Delta_t}{d}, \ldots, \Delta_t$. This perturbation is added to the roll *true* function $\bar{\varphi}$ to produce the *measured* roll function φ. The same is done with the pitch. The resulting noisy attitude coefficients are denoted by Φ.

Then the roll and pitch coefficients are estimated from the noisy ground control points $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ and the noisy attitude coefficients Φ using the proposed refinement algorithm (see Algorithm 3). This results in the *estimated* functions $\hat{\varphi}$ and $\hat{\psi}$.

To assess the algorithm performance, the estimation errors $\hat{\varphi} - \bar{\varphi}$ and $\hat{\psi} - \bar{\psi}$ are compared to the measurement errors $\varphi - \bar{\varphi}$ and $\psi - \bar{\psi}$. In order to facilitate the comparison, the L^2 and L^∞ norms of the error functions are computed. These norms give the root mean square and the maximal errors. The same quantities are computed for the *localization error function* defined by

$$\begin{aligned} [0, \Delta_t] &\to \mathbf{R}_+ \\ t &\mapsto d(L_\Phi(t\delta_t, y_0, h_0), L_{\bar{\Phi}}(t\delta_t, y_0, h_0)), \end{aligned} \tag{42}$$

where y_0 is the principal point coordinate, $h_0 = \frac{1}{n}\sum_i \bar{h}_i$ is the mean altitude and $d(\mathbf{X}, \mathbf{X}')$ is the geodetic distance between two points on the surface of the Earth. As for φ and ψ, the estimation error (computed with $\hat{\Phi}$) is compared to the measurement error (computed with Φ as in (42)).

4.2 Online Demo

The experimental setup presented in Section 4.1 is made available to the reader through the online demo associated to this paper[3].

To help the user with the fastidious task of setting the numerous parameters of the camera, a few predefined sets of parameters are proposed. They correspond to the internal and orbit parameters of well-known commercial satellites such as Pléiades and WorldView and are given in Table 1. The reference node longitude λ_0 and the initial angular position α_0 on the orbit change over time for any satellite, thus they should be chosen by the user of the demo. As these two parameters show no influence on the results of the refinement algorithm, they are not exposed in the demo interface and have fixed values $\lambda_0 = 30°$ and $\alpha_0 = 180°$ for all the experiments.

As it would be difficult to choose manually a set of realistic values for the 12 attitude coefficients, the user only selects an initial pointing direction $\Psi_x, \Psi_y \in [-\frac{\pi}{4}, \frac{\pi}{4}]^2$ and a heading $\gamma \in [0, 2\pi]$. The *heading* defines the desired direction of the pushbroom track on the ground, with respect to the local North direction. The attitude coefficients $\bar{\Phi}$ are then computed from Ψ_x, Ψ_y and γ thanks to the *guidance algorithm*, given in Appendix B.

Next the user selects a number of image points $\bar{\mathbf{x}}_i$ by clicking on a white square representing the image domain. A random altitude $\bar{h}_i$ is associated to each image point.

Finally the user selects the standard deviations σ_{image} and σ_{world} of the Gaussian perturbations added to the coordinates of the image and world points, and selects the degree and amplitude of the polynomial perturbation added to the roll and pitch functions.

As a result of the simulation, the online demo shows the plots of the localization error function (see (42)) as well as the roll and pitch error functions, each computed with the measured data and with the estimated data. An example is given in Figure 9. The complete testing procedure implemented in the online demo is summarized in Algorithm 4.

[3] https://doi.org/10.5201/ipol.2015.146

Algorithm 4: Testing procedure for the roll and pitch refinement algorithm.

Input: $\bar{\delta}_t, \bar{w}, \bar{f}, \bar{y}_0, \bar{a}, \bar{i}$: set of true internal and orbit parameters
$\bar{\lambda}_0, \bar{\alpha}_0$: true reference node longitude and initial angular position on the orbit
Ψ_x, Ψ_y, γ: initial pointing direction and heading
$\{\bar{\mathbf{x}}_k\}$: set of $n \geq 1$ true image locations
$\sigma_{\text{image}}, \sigma_{\text{world}}$: std dev of the Gaussian perturbations added to the correspondences
d, η: degree and amplitude of the perturbation added to the roll and pitch functions

Output: statistics (rmse, max) of the roll, pitch and localization error functions, with measured and estimated attitudes

begin

$\bar{\Phi} \leftarrow$ `guidance`$(\Psi_x, \Psi_y, \gamma, \bar{\delta}_t, \bar{w}, \bar{f}, \bar{y}_0, \bar{a}, \bar{i}, \bar{\lambda}_0, \bar{\alpha}_0)$
for $k \in \{1 \dots n\}$ **do**
  $\bar{h}_k \leftarrow$ random value between 0 and 1000
  $\bar{\mathbf{X}}_k \leftarrow L_{\bar{\Phi}}(\bar{\mathbf{x}}_k, \bar{h}_k)$
for $k \in \{1 \dots n\}$ **do**
  $\mathbf{u} \leftarrow$ random vector on the unit sphere S^2
  $\mathbf{v} \leftarrow$ random vector on the unit circle S^1
  $\mathbf{X}_k \leftarrow \bar{\mathbf{X}}_k + \sigma_{\text{world}}\mathbf{u}$
  $\mathbf{x}_k \leftarrow \bar{\mathbf{x}}_k + \sigma_{\text{image}}\mathbf{v}$
for $\theta \in \{\varphi, \psi\}$ **do**
  $\theta^0, \theta^1, \dots, \theta^d \leftarrow$ uniform random samples in $[-\eta, \eta]$
  $p \leftarrow$ Lagrange polynomial interpolating $(0, \theta^0), (\frac{\Delta_t}{d}, \theta^1), \dots, (\Delta_t, \theta^d)$
  $\theta \leftarrow \bar{\theta} + p$
$\hat{\varphi}, \hat{\psi} \leftarrow$ `roll_and_pitch_refinement`$(\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}, \bar{\delta}_t, \bar{w}, \bar{f}, \bar{y}_0, \bar{a}, \bar{i}, \bar{\lambda}_0, \bar{\alpha}_0, \varphi, \psi, \bar{\omega}, \eta)$
for $\theta \in \{\varphi, \psi\}$ **do**
  return statistics of $\theta - \bar{\theta}$ and $\hat{\theta} - \bar{\theta}$

4.3 Numerical Results

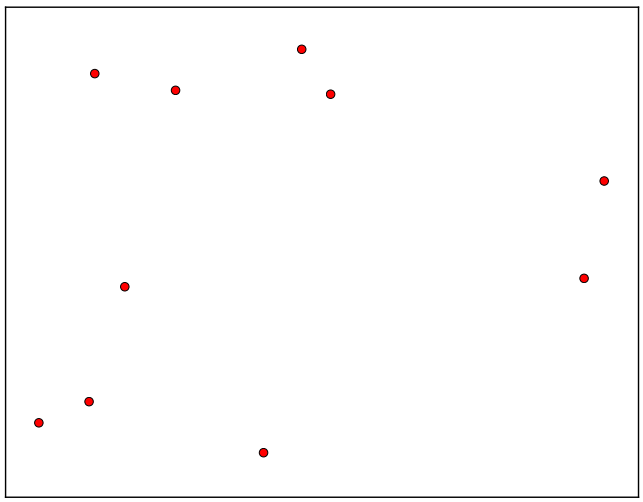
The testing procedure described in Algorithm 4 was run with typical sets of camera parameters on various control points configurations and various roll and pitch errors. Results presented in Table 2 and Figure 10, show that for a given degree d of the roll and pitch polynomial errors, $d+1$ control points are sufficient to reduce the localization error by one order of magnitude.

However this conclusion holds only if the row indices of the control points are roughly allocated regularly over the whole row range. Two control points located on neighboring rows correspond to very close times t and t', thus the two constraints they provide for the roll and pitch fitting behave as only one constraint. This is illustrated by the results in Table 3 and Figure 11.

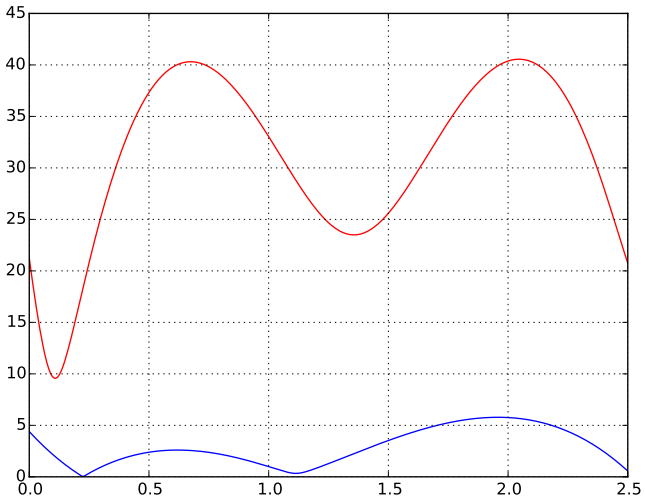
Experiments with different amounts of noise added to the control points show that, as expected, the estimation is perfect in the absence of noise. The accuracy of the estimation reduces as soon as the noise amount increases, but this can be compensated by increasing the quantity of control points. These observations are illustrated by results presented in Table 4 and Figure 12.

5 Conclusion and Future Work

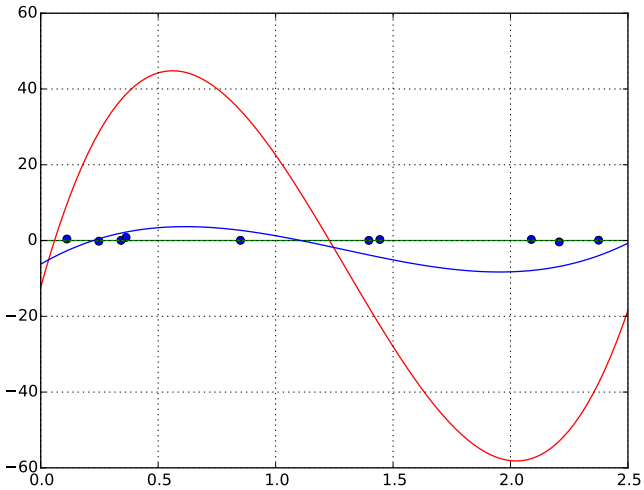
We gave a detailed description of a simplified orbiting pushbroom camera model and showed that, under reasonable assumptions, the main factor responsible for the localization error are the roll



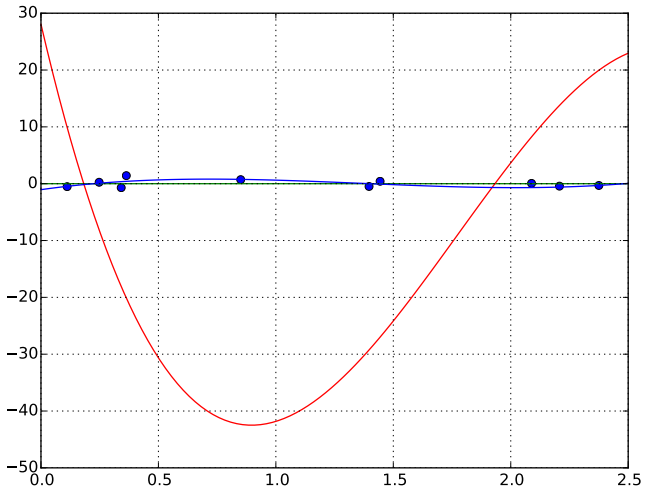
(a) Input image points



(b) Localization error in meters



(c) Roll error in micro radians



(d) Pitch error in micro radians

Figure 9: Example of results produced with the online demo associated to this paper. Red curves show the error before correction and blue curves after correction. The user clicked on the image domain to select the input image points shown in (a). The localization error (in meters) as a function of time (in seconds) is shown in (b). The red curve gives the error made using the measured attitudes, while the blue curve gives the error made using the estimated attitudes. The roll and pitch errors (in micro radians) as functions of time (in seconds) are shown on (c) and (d). The blue dots give the errors of the instantaneous roll and pitch samples computed from each world to image correspondence.

and pitch functions. We presented a simple algorithm to refine the roll and pitch functions using ground control points. It exploits the fact that each ground control point imposes two independent constraints on the instantaneous values of the roll and pitch, thus permits to compute these values. Experimental results show that as soon as one has $n+1$ control points on roughly well spaced rows, a polynomial error of degree no greater than n is substantially corrected by the proposed algorithm. In contrast to classical minimization approaches, this algorithm has the particularity that outliers are automatically detected and removed, preventing them to bias the final result.

To further improve the localization accuracy, once the roll and pitch have been refined, one could refine the yaw coefficients by minimizing the mean quadratic distance between world points and

parameters					results					
	noise		perturbation		roll error (µrad)		pitch error (µrad)		loc. error (m)	
satellite type	σ_{image}	σ_{world}	d	η (µrad)	before	after	before	after	before	after
Pléiades	0.5 px	0.2 m	0	50	76.48	0.01	61.33	0.92	43.03	0.40
Pléiades	0.5 px	0.2 m	1	50	40.11	0.08	39.51	0.59	24.75	0.26
Pléiades	0.5 px	0.2 m	2	50	60.53	1.01	27.42	0.89	29.16	0.59
Pléiades	0.5 px	0.2 m	3	50	34.74	0.94	52.42	5.24	27.62	2.36

Table 2: Numerical parameters and results of the experiments shown in Figure 10, just below. For each polynomial error degree d, only $d+1$ control points are used.

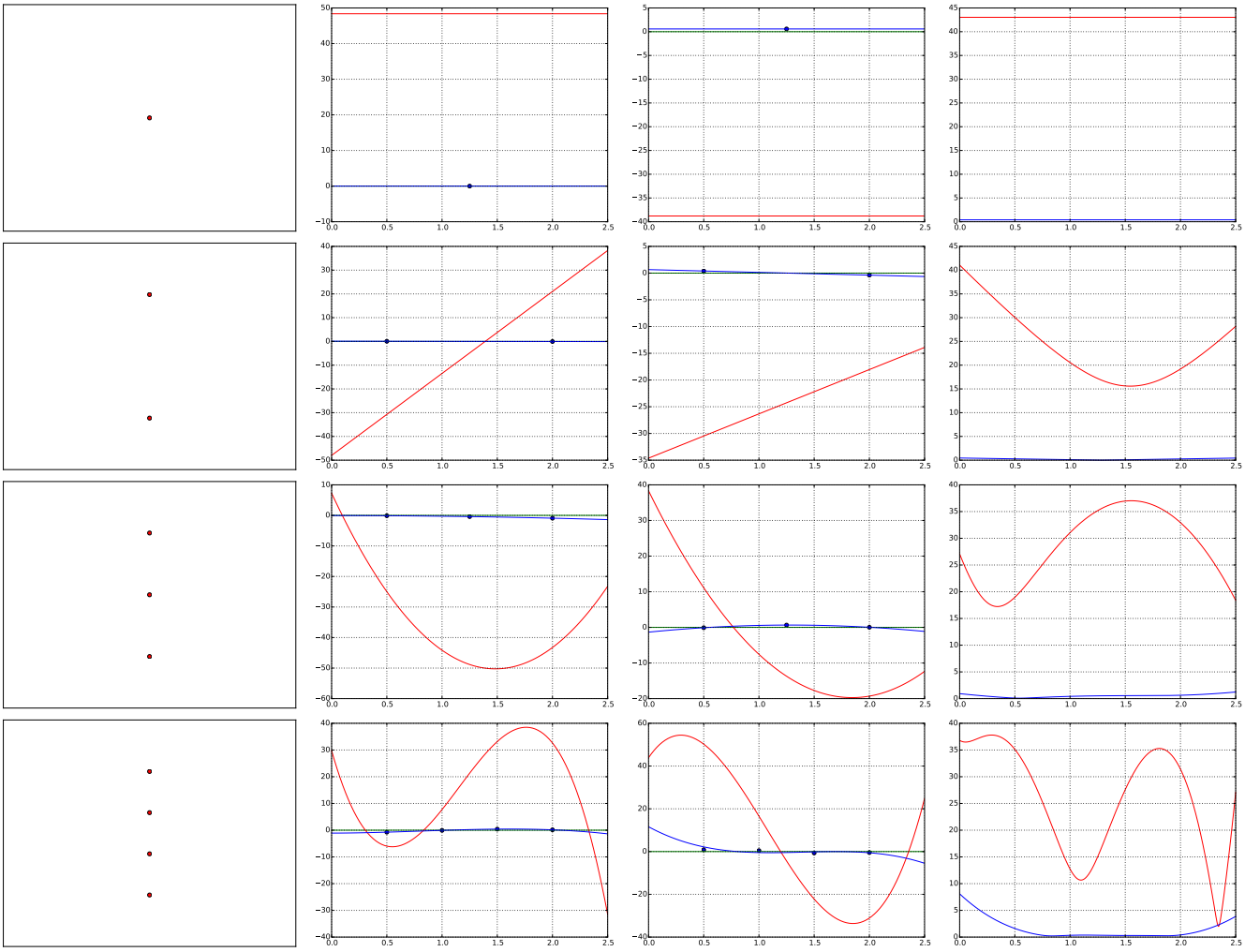


Figure 10: Minimal number of ground control points needed to correct the roll and pitch error, depending on the degree of the polynomial error. Each of the four rows presents a different experiment. On the d^{th} row, a polynomial error of degree $d-1$ is corrected with d control points. The numerical parameters and results of these experiments are given in Table 2.

parameters					results					
	noise		perturbation		roll error (µrad)		pitch error (µrad)		loc. error (m)	
satellite type	σ_{image}	σ_{world}	d	η (µrad)	before	after	before	after	before	after
Pléiades	0.5 px	0.2 m	3	50	10.97	2.96	51.50	2.15	23.11	1.62
Pléiades	0.5 px	0.2 m	3	50	10.97	5.75	51.50	9.58	23.11	4.94
Pléiades	0.5 px	0.2 m	3	50	10.97	18.10	51.50	9.07	23.11	8.89
Pléiades	0.5 px	0.2 m	3	50	10.97	34.91	51.50	58.08	23.11	29.84

Table 3: Numerical parameters and results of the experiments shown in Figure 11, just below. As soon as two control points are located barely on the same row, they count as only one point. This affects the accuracy of the refinement.

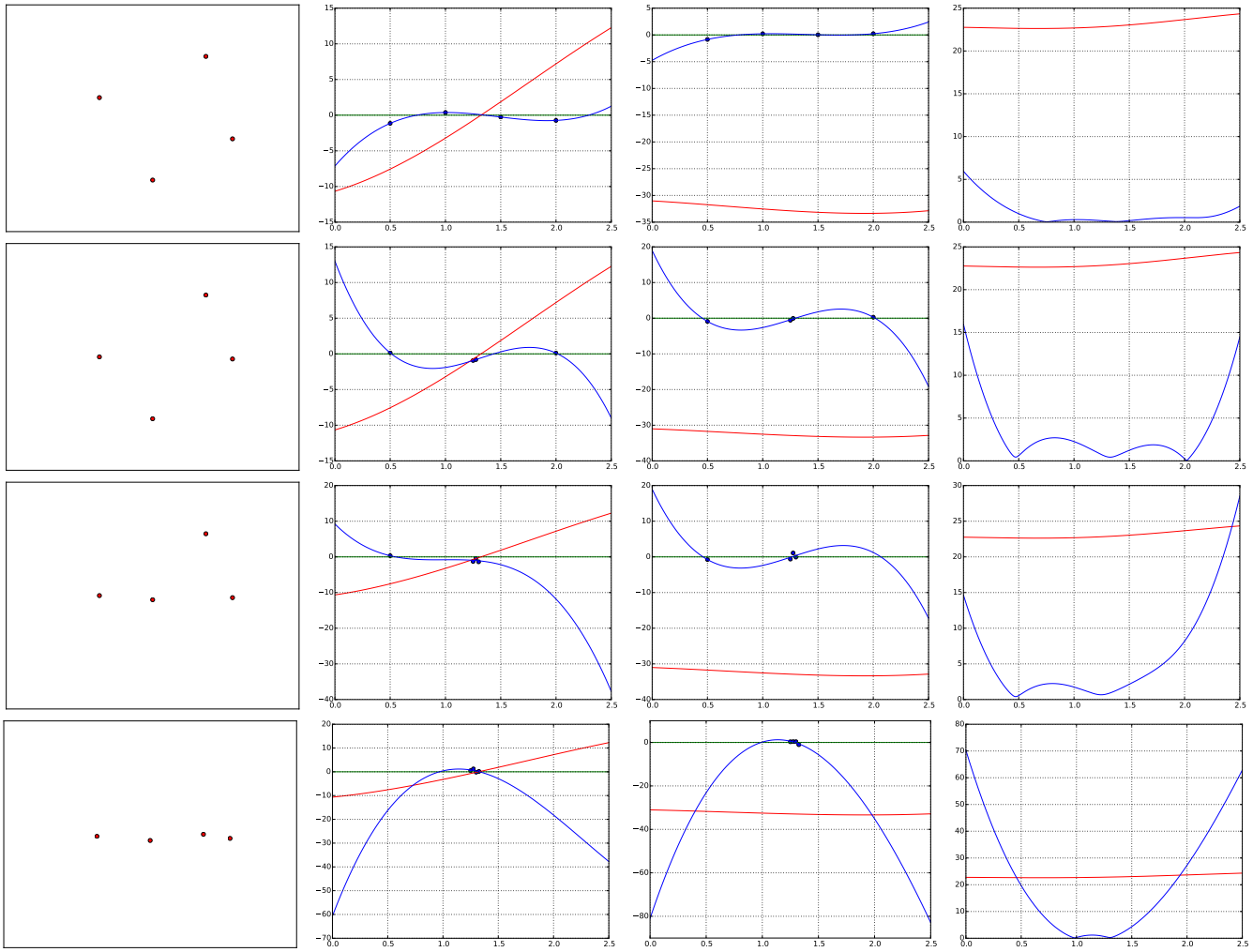


Figure 11: Each of the four rows presents a different experiment. The polynomial error of degree 3 added to the attitudes is identical in the four experiments, as well as the number of control points, but the positions of the control points vary. When two control points are located on neighboring rows in the image plane, their effect on the estimated roll and pitch functions is that of a single control point. This is why on the second experiment shown here the polynomial fitting of the roll and pitch corrections is poorly conditioned. We are trying to fit a polynomial of degree up to 3 to three points. Hence the refinement is less accurate. The situation gets worse on the third and fourth experiments, as confirmed by the numerical results given in Table 3.

parameters					results					
	noise		perturbation		roll error (µrad)		pitch error (µrad)		loc. error (m)	
satellite type	σ_{image}	σ_{world}	d	η (µrad)	before	after	before	after	before	after
Pléiades	0.5 px	0.2 m	3	50	45.87	0.93	31.18	1.01	24.33	0.60
Pléiades	1.0 px	1.0 m	3	50	45.87	8.84	31.18	7.41	24.33	5.07
Pléiades	1.0 px	1.0 m	3	50	45.87	2.72	31.18	2.12	24.33	1.52
Pléiades	2.0 px	2.0 m	3	50	45.87	3.30	31.18	7.50	24.33	3.60
Pléiades	2.0 px	2.0 m	3	50	45.87	1.43	31.18	2.41	24.33	1.23

Table 4: Numerical parameters and results of the experiments shown in Figure 12, just below. As soon as the accuracy of the ground control points decreases, the overall accuracy of the refined attitudes decreases. This effect can be compensated by adding more ground control points.

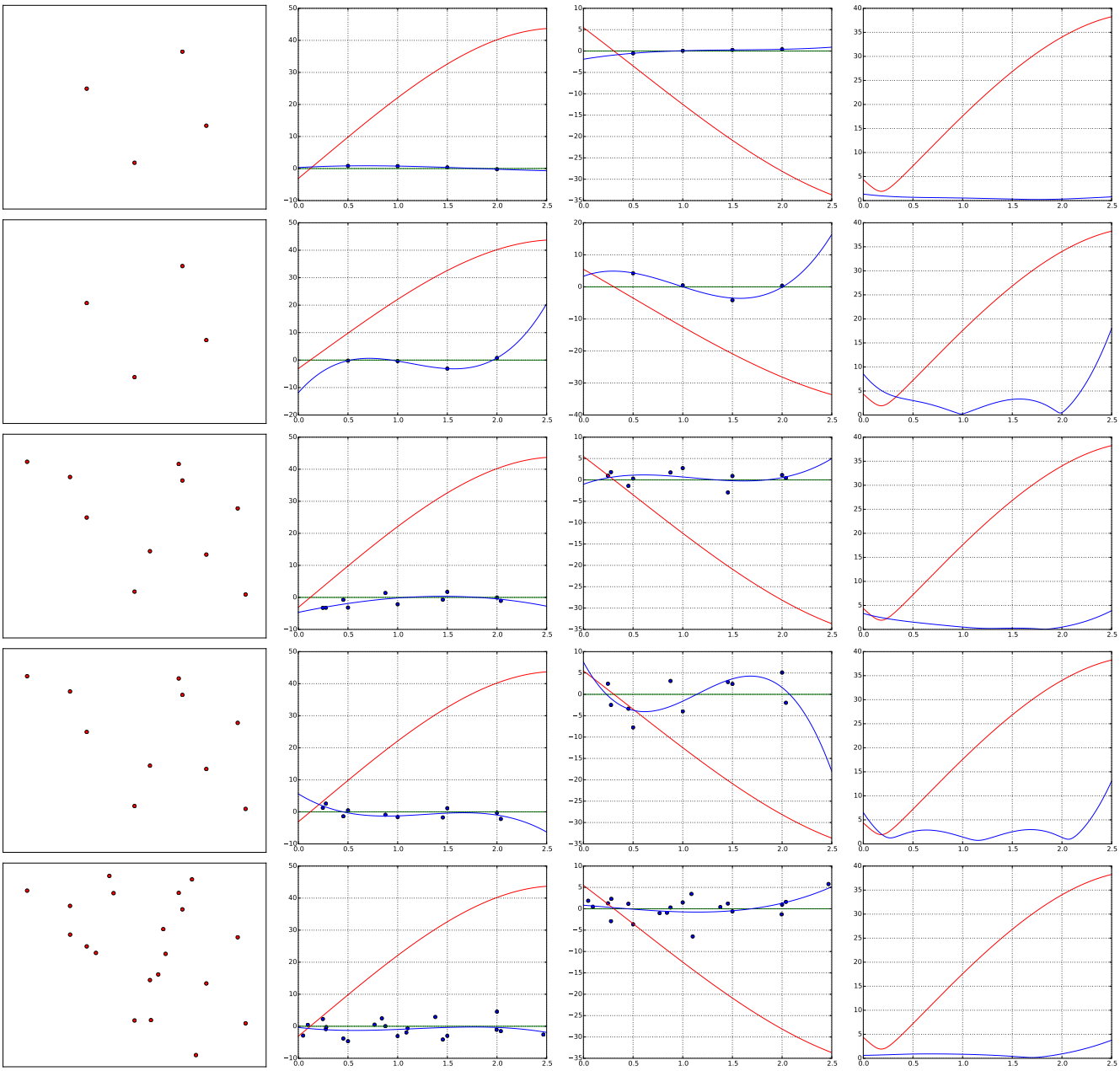


Figure 12: In the first experiment (first row) the roll and pitch are refined from four control points. In the second row, the same experiment is done with a higher level of noise. The accuracy of the output roll and pitch functions is lower. In the third row the same experiment is done with six additional control points. The output is more accurate. Similarly, on the fourth row the noise level is increased and in the fifth row 10 points are added.

back-projected rays (see energy (21)). Other parameters such as the orbit parameters and the focal length may be incorporated in that energy to be refined too. The Hessian of the energy could be used to evaluate the quality of the extrema.

Future work will also include careful experimentation of the proposed algorithm with real data, and study whether it is possible or not to reduce the need for ground control points by using stereo images, i.e. sequences of at least two views of the same area. Refining the orientation parameters of several cameras given a sequence of views of the same scene is a problem known in computer vision and photogrammetry as *bundle adjustment*. Several open source tools have been developed to solve this problem, such as Apero [27], Bundler [34] and OpenMVG [25]. It is likely that ground control points will always be needed because, as was shown by Jama et al. [17], the bundle adjustment problem with pushbroom cameras has several degrees of freedom, even if the camera positions are known.

A Notations

As far as possible we try to use the same notations as Hartley and Zisserman in their reference book on multiple view geometry [13]. Table 5 summarizes them.

B Guidance Algorithm

In our implementation of the pushbroom camera geometric simulator, the true attitude coefficients $\bar{\Phi}$ are not set directly by the user. Instead, the user chooses an initial pointing direction Ψ_x, Ψ_y and a heading γ which gives the direction of the movement of the projection of the pushbroom sensor on the ground, with respect to the local North. The attitude coefficients are computed from Ψ_x, Ψ_y and γ thanks to Algorithm 5.

Acknowledgment

This work was partly founded by Centre National d'Etudes Spatiales (CNES, MISS Project), European Research Council (advanced grant Twelve Labours), Office of Naval research (ONR grant N00014-14-1-0023), DGA Stéréo project, ANR-DGA (project ANR-12-ASTR-0035) and Institut Universitaire de France.

References

[1] E. D. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, 95 (2003), pp. 249–277. http://dx.doi.org/10.1007/s10107-002-0349-3.

[2] E. Berthier, C. Vincent, E. Magnússon, A. P. Gunnlaugsson, P. Pitte, E. Le Meur, M. Masiokas, L. Ruiz, F. Palsson, J.M.C. Belart, and P. Wagnon, *Glacier topography and elevation changes derived from Pléiades sub-meter stereo images*, The Cryosphere, (2014), pp. 2275–2291. http://dx.doi.org/10.5194/tc-8-2275-2014.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004. http://dx.doi.org/10.1017/CBO9780511804441.

Algorithm 5: How to compute attitude coefficients from a given heading and initial pointing.

Input: Ψ_x, Ψ_y: initial pointing direction
γ: heading
h_0: rough altitude of the scene
Δ_t, δ_t: duration of the acquisition and dwell time
a, i, λ_0: orbit parameters
α_0: satellite initial position on the orbit
n: number of samples used to estimated the attitude polynomials
d: degree of the polynomials used to describe the attitude functions

Output: $\varphi, \psi, \omega : [0, \Delta_t] \to \mathbf{R}^3$: attitude polynomial functions

begin

$\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\top \leftarrow$ `target_point_orbital_frame`(Ψ_x, Ψ_y, h_0) *// algorithm 6*

for $k \in \{0, \dots, n-1\}$ **do**

  $\mathbf{v} \leftarrow$ `pushbroom_direction_on_the_ground`$(\mathbf{X}, \gamma, \frac{k\Delta_t}{n-1})$ *// algorithm 7*

  $\varphi^k, \psi^k, \omega^k \leftarrow$ `attitude_from_point_and_speed`$(\mathbf{X}, \mathbf{v})$ *// algorithm 8*

  $\mathbf{X} \leftarrow \mathbf{X} +$ `pixel_projection_width`$(\mathtt{R}(\varphi, \psi, \omega))\frac{\Delta_t}{(n-1)\delta_t}\mathbf{v}$ *// algorithm 9*

  $\mathrm{Z} \leftarrow \mathrm{Z} - (R_E + a)$

  $\mathbf{X} \leftarrow \mathtt{P}_{\text{rotational, orbital}}(\frac{(k+1)\Delta_t}{n-1}, \lambda_0, i, \alpha_0)^{-1}\mathtt{P}_{\text{rotational, orbital}}(\frac{k\Delta_t}{n-1}, \lambda_0, i, \alpha_0)\mathbf{X}$

  $\mathrm{Z} \leftarrow \mathrm{Z} + (R_E + a)$

for $\theta \in \{\varphi, \psi, \omega\}$ **do**

  $\theta \leftarrow$ least squares degree d polynomial fit of points $\{\frac{k\Delta_t}{n-1}, \theta^k\}_{k=0,\dots,n-1}$

return φ, ψ, ω

Algorithm 6: Orbital coordinates of the intersection of the optical axis and the ground

Input: Ψ_x, Ψ_y: initial pointing direction
h_0: altitude of the scene
a: flying altitude of the satellite

Output: $\mathbf{X}$: coordinates of the intersection of the optical axis and the ground

begin

$\mathbf{E} = (0, 0, R_E + a)^\top$

$\mathbf{v} = (\tan(\Psi_y), -\tan(\Psi_x), 1)^\top$

$\mathbf{X} \leftarrow$ intersection of the line directed by $\mathbf{v}$ and the sphere with center $\mathbf{E}$ and radius $R_E + h_0$

Algorithm 7: Pushbroom direction on the ground

Input: $\mathbf{X}$: orbital coordinates of the ground point currently targeted
γ: heading, i.e. direction of the pushbroom array movement, on the ground
t: time elapsed since the beginning of the acquisition, in seconds

Output: $\mathbf{v}$: orbital coordinates of a vector giving the pushbroom movement direction, projected on the ground

begin

$\mathbf{E} \leftarrow (0, 0, R_E + a)^\top$

$\lambda, \theta \leftarrow$ longitude, latitude of $\mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0)(\mathbf{X} - \mathbf{E})$

$\mathbf{v} \leftarrow (0, \sin\gamma, \cos\gamma)^\top$

return $\mathtt{P}_{\text{rotational, orbital}}(t, \lambda_0, i, \alpha_0) \begin{bmatrix} c_\lambda & -s_\lambda & 0 \\ s_\lambda & c_\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_\theta & 0 & -s_\theta \\ 0 & 1 & 0 \\ s_\theta & 0 & c_\theta \end{bmatrix} \mathbf{v}$

Algorithm 8: Compute the satellite attitude from a targeted point and a pushbroom direction

Input: $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z})$: point on Earth, given by its coordinates in the orbital frame
$\mathbf{v}$: movement of the pushbroom array, projected on the ground, given in the orbital frame.

Output: φ, ψ, ω: attitude angles

begin

$\varphi \leftarrow -\arctan(\frac{\mathrm{Y}}{\mathrm{Z}})$

$\psi \leftarrow \arcsin(\frac{\mathrm{X}}{\|\mathbf{X}\|})$

$\mathbf{n} \leftarrow \mathbf{X} - (0, 0, R_E + a)^\top$

$\mathbf{w} = \mathbf{n} \times \mathbf{v}$

$\mathbf{X} = -\mathbf{n} \times \mathtt{R}(\varphi, \psi, 0)(0, 1, 0)^\top$

$\mathbf{Y} = -\mathbf{n} \times \mathtt{R}(\varphi, \psi, 0)(1, 0, 0)^\top$

if $\mathbf{w} \cdot \mathbf{Y} = 0$ **then**

  $\omega \leftarrow \frac{\pi}{2}$

else

  $\omega \leftarrow \arctan(\frac{\mathbf{w} \cdot \mathbf{X}}{\mathbf{w} \cdot \mathbf{Y}})$

if $\mathbf{v} \cdot \mathbf{Y} \sin\omega < \mathbf{v} \cdot \mathbf{X} \cos\omega$ **then**

  if $\omega < 0$ **then**

    $\omega \leftarrow \omega + \pi$

  else

    $\omega \leftarrow \omega - \pi$

return φ, ψ, ω

Algorithm 9: Compute the width of the projection of a pixel on the ground

Input: $\mathtt{M}$: change of coordinates matrix between orbital frame and camera frame
$\mathbf{x} = (x, y)$: coordinates of the pixel in the image plane

Output: width, in meters, of the projection of the pixel on the ground, in the pushbroom direction (orthogonal to the pushbroom movement)

begin

1. $\mathbf{v} \leftarrow \mathtt{M}(-x, -y, f)^\top$
2. $\mathbf{E} \leftarrow (0, 0, R_E + a)^\top$
3. $\mathbf{X} \leftarrow$ intersection of the line directed by $\mathbf{v}$ and the sphere with center $\mathbf{E}$ and radius R_E
4. $\Pi \leftarrow$ plane normal to $\mathbf{X} - \mathbf{E}$ passing by $\mathbf{X}$
5. $\mathbf{v}_1 \leftarrow \mathtt{M}(-x - \frac{w}{2}, -y, f)^\top$
6. $\mathbf{v}_2 \leftarrow \mathtt{M}(-x + \frac{w}{2}, -y, f)^\top$
7. $\mathbf{p}_1 \leftarrow$ intersection of the line directed by $\mathbf{v}_1$ with the plane Π
8. $\mathbf{p}_2 \leftarrow$ intersection of the line directed by $\mathbf{v}_2$ with the plane Π
9. return $\|\mathbf{p}_1 - \mathbf{p}_2\|$

[4] J.B. Case, *The analytical reduction of panoramic and strip photography*, Photogrammetria, 22 (1967), pp. 127–141. http://dx.doi.org/10.1016/0031-8663(67)90026-9.

[5] M. Crespi, F. Fratarcangeli, F. Giannone, and F. Pieralice, *SISAR: a rigorous orientation model for synchronous and asynchronous pushbroom sensors imagery*, in International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (1/W51), Hannover, 2007. http://www.isprs.org/proceedings/XXXVI/1-W51/.

[6] ——, *High Resolution Satellite Image Orientation Models*, in Geospatial Technology for Earth Observation, Springer US, Boston, MA, 2010, ch. 4, pp. 63–104. http://dx.doi.org/10.1007/978-1-4419-0050-0_4.

[7] F. de Lussy, D. Greslou, C. Dechoz, V. Amberg, J.-M. Delvit, L. Lebegue, G. Blanchet, and S. Fourest, *Pléiades HR in flight geometrical calibration: location and mapping of the focal plane*, in ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIX-B1, jul 2012, pp. 519–523. http://dx.doi.org/10.5194/isprsarchives-XXXIX-B1-519-2012.

[8] D. Fritsch and D. Stallmann, *Rigorous Photogrammetric Processing of High Resolution Satellite Imagery*, in ISPRS Archives, vol. XXXIII, Amsterdam, 2000.

[9] D.J. Gugan, *Practical aspects of topographic mapping from SPOT imagery*, The Photogrammetric Record, 12 (1987), pp. 349–355. http://dx.doi.org/10.1111/j.1477-9730.1987.tb00581.x.

[10] H. Guichard, *Etude théorique de la précision dans l'exploitation cartographique d'un satellite à défilement. Application à SPOT*, Bulletin - Société Française de Photogrammétrie et de Télédétection, (1983), pp. 15–26. http://cat.inist.fr/?aModele=afficheN&cpsidt=9463022.

[11] R. Gupta and R. Hartley, *Camera Estimation for Orbiting Pushbrooms*, in Asian Conference on Computer Vision, 1995.

[12] ——, *Linear pushbroom cameras*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997), pp. 963–975. http://dx.doi.org/10.1109/34.615446.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second ed., apr 2004. http://dx.doi.org/10.1017/CBO9780511811685.

[14] K. Jacobsen, *Geometric Calibration of Space Remote Sensing Cameras for Efficient Processing*, in ISPRS Archives, 1998, pp. 33–43.

[15] ——, *Accuracy of Digital Orthophotos From High Resolution Space Imagery*, in Proceedings of ISPRS Workshop on High Resolution Mapping from Space, Hannover, 2003.

[16] ——, *DEM Generation by SPOT HRS*, in ISPRS Archives – Volume XXXV Part B1, Orhan Altan, ed., Istanbul, 2004, pp. 439–444. http://www.isprs.org/proceedings/XXXV/congress/comm1/comm1.aspx.

[17] M. Jama, C. Lewis, and D.E. Schinstock, *Identifying degrees of freedom in pushbroom bundle adjustment*, ISPRS Journal of Photogrammetry and Remote Sensing, 66 (2011), pp. 400–407. http://dx.doi.org/10.1016/j.isprsjprs.2011.02.001.

[18] G. Joseph, *Building Earth Observation Cameras*, CRC Press, jan 2015. http://dx.doi.org/10.1201/b18022.

[19] G. Konecny, *Geoinformation*, Taylor & Francis, Abingdon, UK, second ed., 2003. ISBN 9781420068566.

[20] G. Konecny, P. Lohmann, H. Engel, and E. Kruck, *Evaluation of SPOT Imagery on Analytical Photogrammetric Instruments*, Photogrammetric Engineering & Remote Sensing, 53 (1987), pp. 1223–1230. http://eserv.asprs.org/PERS/1987journal/sep/1987_sep_1223-1230.pdf.

[21] V. Kratky, *Universal Photogrammetric Approach to Geometric Processing of SPOT Images*, in XVIth ISPRS Congress Technical Commission IV: Cartographic and Data Bank Applications of Photogrammetry and Remote Sensing, Kyoto, 1988, ISPRS Archives, pp. 180–189. http://www.isprs.org/proceedings/xxvii/congress/part4/.

[22] V Kratky, *On-Line Aspects of Stereophotogrammetric Processing of SPOT Images*, Photogrammetric Engineering & Remote Sensing, 55 (1989), pp. 311–316. http://eserv.asprs.org/PERS/1989journal/mar/1989_mar_311-316.pdf.

[23] V. Kratky, *Rigorous photogrammetric processing of SPOT images at CCM Canada*, ISPRS Journal of Photogrammetry and Remote Sensing, 44 (1989), pp. 53–71. http://dx.doi.org/10.1016/0924-2716(89)90007-5.

[24] P. Michalis and I. Dowman, *A Generic Model for Along-track Stereo Sensors Using Rigorous Orbit Mechanics*, Photogrammetric Engineering & Remote Sensing, 74 (2008), pp. 303–309. http://dx.doi.org/10.14358/PERS.74.3.303.

[25] P. Moulon, P. Monasse, and R. Marlet, *Adaptive Structure from Motion with a Contrario Model Estimation*, in Asian Conference on Computer Vision, Daejeon, 2013, Springer, pp. 257–270. http://dx.doi.org/10.1007/978-3-642-37447-0_20.

[26] R. Perrier, E. Arnaud, P. Sturm, and M. Ortner, *Estimation of an Observation Satellite's Attitude using Multimodal pushbroom Cameras*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 37 (2015), pp. 987–1000. http://dx.doi.org/10.1109/TPAMI.2014.2360394.

[27] M. Pierrot Deseilligny and I. Clery, *Apero, an Open Source Bundle Adjusment Software for Automatic Calibration and Orientation of Set of Images*, in ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII-5, 2012, pp. 269–276. http://dx.doi.org/10.5194/isprsarchives-XXXVIII-5-W16-269-2011.

[28] D. Poli, *Modelling of spaceborne linear array sensors*, PhD thesis, Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, ETH, Zurich, 2005. http://dx.doi.org/10.3929/ethz-a-005051521.

[29] ——, *A Rigorous Model for Spaceborne Linear Array Sensors*, Photogrammetric Engineering & Remote Sensing, 73 (2007), pp. 187–196. http://dx.doi.org/10.14358/PERS.73.2.187.

[30] D. Poli, F. Remondino, E. Angiuli, and G. Agugiaro, *Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction*, ISPRS Journal of Photogrammetry and Remote Sensing, 100 (2014), pp. 35–47. http://dx.doi.org/10.1016/j.isprsjprs.2014.04.007.

[31] D. Poli and T. Toutin, *Review of developments in geometric modelling for high resolution satellite pushbroom sensors*, The Photogrammetric Record, 27 (2012), pp. 58–73. http://dx.doi.org/10.1111/j.1477-9730.2011.00665.x.

[32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, third ed., 2007. http://dx.doi.org/10.2307/1269484.

[33] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill Education, third ed., 1976. http://shop.mheducation.com/highered/product.M007054235X.html.

[34] N. Snavely, S.M. Seitz, and R. Szeliski, *Photo tourism: Exploring Photo Collections in 3D*, ACM Transactions on Graphics, 25 (2006), pp. 835–846. http://dx.doi.org/10.1145/1141911.1141964.

[35] T. Toutin, *Analyse mathématique des possibilités cartographiques du satellite SPOT*, PhD thesis, Ecole Nationale des Sciences Géographiques (ENSG), Champs-sur-Marne, France, 1985.

[36] ——, *Error tracking in IKONOS geometric processing using a 3D parametric model*, Photogrammetric Engineering & Remote Sensing, (2003), pp. 1–32. http://www.asprs.org/a/publications/pers/2003journal/january/2003_jan_43-51.pdf.

[37] ——, *Comparison of stereo-extracted DTM from different high-resolution sensors: SPOT-5, EROS-A, IKONOS-II, and QuickBird*, IEEE Transactions on Geoscience and Remote Sensing, 42 (2004), pp. 2121–2129. http://dx.doi.org/10.1109/TGRS.2004.834641.

[38] ——, *Review article: Geometric processing of remote sensing images: models, algorithms and methods*, International Journal of Remote Sensing, 25 (2004), pp. 1893–1924. http://dx.doi.org/10.1080/0143116031000101611.

[39] ——, *Generation of DSMs from SPOT-5 in-track HRS and across-track HRG stereo data using spatiotriangulation and autocalibration*, ISPRS Journal of Photogrammetry and Remote Sensing, 60 (2006), pp. 170–181. http://dx.doi.org/10.1016/j.isprsjprs.2006.02.003.

[40] ——, *Spatiotriangulation with multisensor HR stereo-images*, IEEE Transactions on Geoscience and Remote Sensing, 44 (2006), pp. 456–462. http://dx.doi.org/10.1109/TGRS.2005.861005.

[41] L. Vandenberghe, *The CVXOPT linear and quadratic cone program solvers*, tech. report, UCLA, 2010. http://www.seas.ucla.edu/~vandenbe/publications/coneprog.pdf.

[42] T. Westin, *Precision Rectification of SPOT Imagery*, Photogrammetric Engineering & Remote Sensing, 56 (1990), pp. 247–253. http://eserv.asprs.org/PERS/1990journal/feb/1990_feb_247-253.pdf.

[43] H. Yésou, A. Escudier, S. Battiston, J.-Y. Dardillac, S. Clandillon, C. Uribe, M. Caspard, H. Giraud, J. Maxant, A. Durand, K. Fellah, M. Studer, C. Huber, L. Philippoteaux, P. de Fraipont, and D. Fontannaz, *Exploitation de l'imagerie Pléiades en cartographie réactive suite à des catastrophes naturelles ayant affecté le territoire français en 2013*, Revue Française de Photogrammétrie et de Télédétection, 209 (2015), pp. 39–45. http://www.sfpt.fr/rfpt/index.php/RFPT/article/view/210.

Notation	Meaning
t	time in seconds from the beginning of the acquisition of the current image
$\mathbf{x}$	image point, usually given by its row and column coordinates
x, y	row and column coordinates of an image point
$\mathbf{X}$	3-space point, given either by its geographic or Cartesian coordinates
$\mathrm{X}, \mathrm{Y}, \mathrm{Z}$	Cartesian coordinates of a 3-space point
λ, θ	longitude, latitude of a point in the Earth-centered rotational frame
h	height of a point above the WGS 84 reference ellipsoid
f	focal length
y_0	principal point coordinate in the sensor frame
w	pixel width
$\varphi(t)$	sensor roll angle at time t measured in the local orbital frame
$\psi(t)$	sensor pitch angle at time t measured in the local orbital frame
$\omega(t)$	sensor yaw angle at time t measured in the local orbital frame
$\mathtt{R}(\varphi, \psi, \omega)$	3D rotation defined by intrinsic roll, pitch and yaw angles
Φ	set of attitude coefficients $\{(\varphi_k, \psi_k, \omega_k)\}_{k=0,\ldots,3}$
Ψ_x, Ψ_y	initial direction of the optical axis in the local orbital frame
γ	direction of the projection of the pushbroom sensor on the ground, w.r.t. the North
$\mathbf{u}$	direction of a back-projected ray from an image point
a	flying altitude of the satellite above the Earth surface
i	inclination of the orbital plane
λ_0	longitude of the reference node
α_0	initial angular position of the satellite on the orbit
α_t	current angular position of the satellite
Δ_t	duration of the acquisition
T_S	orbital period of the satellite
R_E	radius of the Earth (assumed to be spherical)
μ	standard gravitational parameter of the Earth
T_E	duration of an Earth stellar day
$\mathbf{E}$	Earth's center of mass
$\mathbf{S}$	satellite center of mass
L	localization function, converting image to geographic coordinates
P	projection function, converting geographic to image coordinates
P_0	initial estimation of the projection function
$J_{x,y}L$	Jacobian of L with respect to x and y
D	swath width of the satellite
n	number of ground control points
d	degree of a polynomial function
σ_{image}	standard deviation of the zero-mean Gaussian noise added to the image coordinates
σ_{world}	standard deviation of the zero-mean Gaussian noise added to the ground coordinates
η	maximum of the polynomial error functions added to the true attitudes

Table 5: Notations used in the paper