Satellite Image Registration for Attitude Estimation with a Constrained Polynomial Model
Régis Perrier, Élise Arnaud, Peter Sturm, Mathias Ortner

To cite this version:

HAL Id: inria-00514896
https://inria.hal.science/inria-00514896
Submitted on 3 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Satellite image registration has been investigated for several years. Nevertheless, little attention has been paid to the linear geometry of the satellite’s imaging sensor, often consisting of several pushbroom cameras. Each pushbroom camera captures 1-D images and uses straight motion of the satellite to build a 2-D image. Yet, attitude variations of the satellite during the acquisition process can lead to significant distortions in the 2-D image. In this paper, we expose the problem and present a constrained image registration method to estimate the satellite’s attitude variations, and thus correct the distorted images. We use a Lucas Kanade framework and a piecewise polynomial model under constraints to deduce the registration equation. The performances of our algorithm are shown on two satellite datasets.

Index Terms— satellite, image registration, pushbroom camera, piecewise polynomial model

1. INTRODUCTION
Satellite image processing has been widely investigated over the last decades in several applications: image registration, super-resolution, image fusion and 3-D reconstruction [1]. In most cases, those algorithms do not take care of the geometry of the camera, considering images as being taken by a classical pinhole camera. Yet in remote sensing, pushbroom cameras are preferred over pinhole cameras for their ability to provide high resolution images at a lower cost [2].

The pushbroom camera is a linear sensor; mounted on a platform which moves orthogonally to the camera axis (as can be seen in figure 1(a)), the sensor sweeps out a region of space taking 1-D images at regular time instants. Stiching them gives a complete 2-D image of the observed scene. During the acquisition process of the 1-D images, the stability of the imaging platform is crucial to build a distortion-free 2-D image. Until now, the attitude of the satellite was assumed to be constant in most applications [2, 3]. However, the recent manufacture of smaller satellites with higher sampling resolution has weakened this assumption resulting in possible warps in images. Figure 1(b) shows a synthetic example of deformations which can occur when the pushbroom camera is tilting around its 3 rotation axes.

Fig. 1. (a) Pushbroom acquisition principle: the camera is moving straight along the x axis and recording 1-D images over time denoted by n; y is the camera axis and z the orthogonal axis to the image plane. We define the orientation of the camera with the yaw (rotation about z), the roll (rotation about x) and the pitch (rotation about y). (b) Example of warps in a regular checkerboard when the pushbroom camera is tilting around its 3 rotation axes.

Though this problem was already known in airborne imagery [4], it has been noticed only recently on satellite images [5]. Indeed, the spacecraft is continuously exposed to space turbulences and vibrations of its engines, making unknown satellite motions more probable. The use of inertial sensors measurements has been suggested to estimate attitude variations and get a distortion-free image [4]. While this can be applied in airborne imagery, space constraints restrict the use of such sensors for cost reasons. Moreover, the sampling rate of inertial sensors (4 to 16Hz) is usually much lower than the sampling rate of each pushbroom camera (around 2500Hz), leading to an information loss on the satellite’s attitude.

Observation satellites have several pushbroom cameras of different radiometric modalities; figure 2 presents a typical satellite focal plane with 3 pushbroom cameras. All images are deformed due to the unknown motions of the focal plane. There is no way that one can retrieve the vibrations using only
one 2-D image by correlating each line with its neighbors. Considering all 2-D images, one can try to register them pair-wise to estimate the deformations. This has been suggested in [5], but their approach uses local correlation on features points which keeps from considering the problem globally. Thus they cannot consider a multi-frame image alignment as described in [6], yielding a sub-optimal estimate. In this paper, we present an original image registration method to recover the attitude of the satellite using only image. We use a direct method, also known as Lucas Kanade based registration algorithm [7], to globally register images in a same coordinate system. We model the attitude variations with a piecewise polynomial function under constraints. This yields a fast and simple algorithm with no regularizing parameter to tune.

\[ T_{\tau_{ij}}(n, m) = \begin{cases} 1 & \text{for } n = m - \tau_{ij}, \text{and } n \not\in [N - \tau_{ij}, N] \\ \cup [2N - \tau_{ij}, 2N] \cup [3N - \tau_{ij}, 3N] & \text{elsewhere} \end{cases} \]

The global registration equation to minimize is:

\[ \hat{\theta} = \arg\min_{\theta} \sum_{i,j: i \neq j} \sum_{y \in S} \left( I_i(W(y; \theta)) - f_{ij}(I_{\tau_{ij}}(W(y; T_{\tau_{ij}} \theta))) \right)^2. \]

As opposed to the classical Lucas Kanade formulation, there is no reference frame here as both images are warped. Considering a reference frame would lead us to the estimation of a relative estimate of the warps that would imply a convolution process to recover the absolute attitude. This is a major drawback of [5]. In our case, we define a method to directly estimate the absolute attitude. Let’s also note that we use a multi-frame registration formulation to estimate \( \theta \), this has been seen to be more efficient for motion estimation [6]. Direct estimation of \( \theta \) in equation (5) is highly ill-posed if no constraints are defined. One could add a prior term to this equation, but the choice of the regularising function is tricky. Moreover, the computation of the regularising parameter which tunes the belief between the image term and the prior term is expensive and has no trivial solution. Going further into the origins of the problem, we know that the attitude variations mainly originate in the engines of the satellite. Such a vibratory process usually has a stationary form in time, so that smoothness is a trustful property of the signal.

### 2.2. Constrained polynomials

To model the vibrations, we use a piecewise polynomial function where polynomials are linked together by constraints.
Polynomials are denoted by $\Phi_{k,\alpha}, \ k \in [1, M]$. $M$ is the number of polynomials needed to fit the attitude variations, and $n_k$ for $k \in [1, M - 1]$ the time instants where constraints are defined between two successive polynomials. This can be related to splines [8], except that our knots are spaced by several time samples as this can be seen on figure 3.

The vibrations can be described with $M$ polynomials:

$$\theta_{\alpha}(n) = \sum_{k=1}^{M} w_k(n) \Phi_{k,\alpha}(n),$$  \hspace{1cm} (6)

where the logical weights $w_k(n)$ are defined such that:

$$w_k(n) = \begin{cases} 1 & \text{for } n \in [n_{k-1}, n_k[ \\ 0 & \text{elsewhere.} \end{cases}$$  \hspace{1cm} (7)

The polynomials of order $P$ are defined as:

$$\Phi_{k,\alpha}(n) = \sum_{p=0}^{P} a_{p,k,\alpha} n^p.$$  \hspace{1cm} (8)

$a_{p, k, \alpha}$ are the polynomial coefficients which fully characterise the attitude variations. We call $a$ the $(3MP \times 1)$ vector which stacks all the coefficients for the attitude estimate. We define the constraints between two polynomials in the same spirit as splines for all $n_k$ with $k \in [1, M - 1]$:

$$\Phi^{(p)}_{k,\alpha}(n_k) = \Phi^{(p)}_{k+1,\alpha}(n_k) \text{ for } p \in [0, P - 1].$$  \hspace{1cm} (9)

This way, continuity is imposed for the derivatives of the polynomial up to order $P - 1$. As for splines, there is only one degree of freedom per time segment $[n_{k-1}, n_k]$ [8]. Notice that equation (9) defines linear equality constraints with $a$. We call $C$ the $(3(M - 1)P \times 3MP)$ sparse matrix such that equation (9) can be written in the following matrix form:

$$Ca = 0.$$  \hspace{1cm} (10)

with $0$ the null vector of size $(3(M - 1)P \times 1)$. Also, concatenating linear equations (6), (7) and (8) yields:

$$\theta = Ha,$$  \hspace{1cm} (11)

where $H$ is a sparse matrix of size $(3N \times 3MP)$ which combines $w_k(n)$ of equation (6) and $n^p$ of equation (8). We can reformulate the image registration problem of equation (5) as a non linear problem with linear equality constraints:

$$\hat{a} = \arg \min_a \sum_{i,j} \sum_{y \in S} (I_i(W(y; Ha)) - f_{ij}(I_{ij}(W(y; T_{\tau ij} Ha))))^2, \text{ such that } Ca = 0.$$  \hspace{1cm} (12)

This global non linear equation is minimised iteratively. The equality constraints are used to reduce the problem to a lower dimensionality with a QR factorisation, then the remaining variables are estimated with a least square procedure.

In practice, the $f_{ij}$ function introduced in section 2 to correct radiometric differences is linear such that:

$$f_{ij}(I_j(y)) = r_{ij,1} + r_{ij,2} I_j(y).$$  \hspace{1cm} (13)

We estimate radiometric coefficients $[r_{ij,1}, r_{ij,2}]$ for each pair of images, and once for all at the beginning of the algorithm with a least square method. The warp function $W$ has a tractable analytical formulation for the three rotation angles if the observed scene is planar. This assumption is approximately true in some cases and the analytical warp shows nice results. In other cases, we use numerical derivatives correlated with a Digital Elevation Model (DEM) data to estimate the attitude variations. We choose the second option in the following results.

3. EXPERIMENTS AND DISCUSSION

We present experimental results on two satellite datasets which have been simulated by ASTRIUM. The ground truth on attitude variations is available, but the simulation process is totally unknown to us. It creates real-life condition data acquisition and can be considered as difficult as could be real data. We use a Matlab implementation on a Core2 duo at 3GHz with 3.8GiB to process the data. In both cases, cameras 1-2, and 2-3 are spaced respectively by 40 and 20 time samples in accordance to the focal plane in figure 2 ($\tau_{12}=40$ and $\tau_{23}=20$). All images are of size $(2564 \times 900)$ pixels.

The first dataset is composed of multispectral data with blue, green and red images. We set the polynomials order $P$ to 3 and their size to 15 time samples. The algorithm took 230 seconds to converge in 10 iterations; the standard deviation of the error compared to ground truth is below $15/100$ either for the roll and the pitch (see fig. 4(a)). The second dataset is composed of monomodal images. This case is considered to be a challenging one as the vibrations contain very low frequency components. We set the polynomial order $P$ to 3 and their size to 20 time samples. The algorithm took 410 seconds to converge in 20 iterations; the standard deviation of the error compared to ground truth is below $20/100$ either for the roll and the pitch (see fig. 4(b)). Let us remark that no estimation of the yaw is given in both experiments. The yaw causes deformations lower than $2/100$ pixel on each vertical side of the images. In fact, the deformations are dominated by geometrical distortions induced by the roll and pitch.
The results on figures 4(a) and 4(b) show performances and some limits of our algorithm. Most of the frequencies are recovered, and the registration process is able to rectify images with an accuracy much below than the pixel unit. Constrained polynomials are suited to the estimation of such vibration processes mainly because of their stationary properties. However in each case, the residual error has a high frequency component. This is mainly linked to the way the polynomial can fit the vibrations in a given time window. For larger time windows, the polynomial will recover low frequency components. For shorter time windows, it has the ability to fit high frequency vibrations but eq. (12) has more degrees of freedom and can become ill-posed. To a lesser extent, the polynomial degree also influences the attitude estimate. Ways to choose the size of the time window, as well as the polynomial degree, are currently investigated.

The originality of our work is to present a global image registration technique for images acquired by pushbroom cameras on board of a satellite. This can solve the problem of attitude variation of the satellite during the process of acquisition. Although in general, the registration problem is highly ill-posed in this case, the algorithm does not need any regularisation due to the built-in realistic assumption of stationarity.

4. REFERENCES