HAL open science

Robust Rational Polynomial Camera Modelling for SAR and Pushbroom Imaging

Roland Akiki, Roger Marí, Carlo de Franchis, Jean-Michel Morel, Gabriele Facciolo

To cite this version:

Roland Akiki, Roger Marí, Carlo de Franchis, Jean-Michel Morel, Gabriele Facciolo. Robust Rational Polynomial Camera Modelling for SAR and Pushbroom Imaging. 2021. hal-03189163

HAL Id: hal-03189163

https://hal.science/hal-03189163v1

Preprint submitted on 2 Apr 2021

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.

arXiv:2102.13423v1 [eess.IV] 26 Feb 2021

Robust Rational Polynomial Camera Modelling for SAR and Pushbroom Imaging

Roland Akiki[1,2] Roger Marí [1] Carlo de Franchis[1,2]
Jean-Michel Morel[1] Gabriele Facciolo[1]

[1] Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France
[2] Kayrros SAS

Abstract

The Rational Polynomial Camera (RPC) model can be used to describe a variety of image acquisition systems in remote sensing, notably optical and Synthetic Aperture Radar (SAR) sensors. RPC functions relate 3D to 2D coordinates and vice versa, regardless of physical sensor specificities, which has made them an essential tool to harness satellite images in a generic way. This article describes a terrain-independent algorithm to accurately derive a RPC model from a set of 3D-2D point correspondences based on a regularized least squares fit. The performance of the method is assessed by varying the point correspondences and the size of the area that they cover. We test the algorithm on SAR and optical data, to derive RPCs from physical sensor models or from other RPC models after composition with corrective functions.

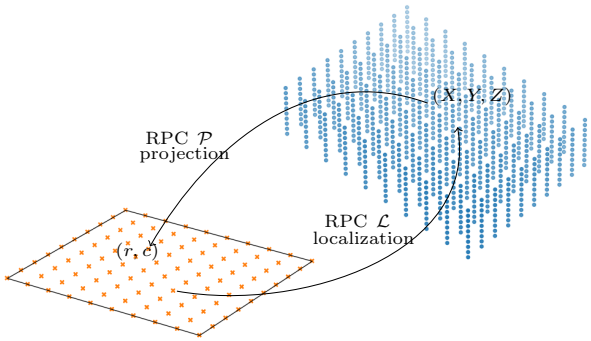

Figure 1: The RPC model is derived using a grid of 3D CNPs (Control Points) and its projection onto the satellite image.

1 Introduction

Developing a remote sensing application requires a set of tools, one of which is geolocation. Geolocation relates the 3D world coordinates to the 2D image. This is represented by means of a *projection* function $\mathcal{P} : \mathbf{R}^3 \to \mathbf{R}^2$, that maps 3D points to the image plane, and its inverse, the *localization* function $\mathcal{L} : \mathbf{R}^2 \times \mathbf{R} \to \mathbf{R}^3$. When all the physical phenomena and components involved in the acquisition process are known, the geolocation functions can be defined by a chain of operations that model such factors, in what is known as a *physical* or *rigorous* sensor model.

Pushbroom scanners are the most common optical satellite image acquisition system, typically consisting of a single line of pixel sensors mounted on a platform that captures each line of the image at a different moment in time. As a result, the exterior orientation parameters, i.e. the perspective center and the attitude angles, change from line to line. The intrinsic parameters (e.g. pixel size, focal length, lens distortion), related to the physical design of the sensor, are constant across the image [1]. A detailed description of a simplified physical sensor model for pushbroom scanners can be found in [2]. In the case of Synthetic Aperture Radar (SAR) images, the most used physical sensor model is the Range-

Doppler model detailed in [3]. SAR satellites send an electromagnetic wave that is reflected on the ground. The image is acquired line by line (similar to a push-broom system), and the position of a ground patch in the image is related to its distance to the sensor, known as the *range*. The Range-Doppler model is constructed based on ephemeris data (time, position and velocity samples along the orbit) and the acquisition timing information. The ephemeris data needs to be interpolated to obtain continuous geolocation functions along the orbit.

Image vendors have adopted the generic Rational Polynomial Camera[1] (RPC) model to save customers from having to deal with the complex specificities of rigorous sensor models. The RPC model is independent of physical properties and offers flexibility to work with different coordinate systems. RPCs have become essential metadata to process satellite images in a generic way, from different sources and for multiple tasks, e.g. photogrammetry and radargrammetry based 3D reconstruction or image ortho-rectification and coregistration.

In this article, we describe a terrain-independent algorithm to fit a RPC model from a physical sensor model or any other geolocation model. Our contributions are:

- An open-source implementation of the method as an easy-to-use Python package, which is available at `https://github.com/cmla/rpcfit`.
- An evaluation of the algorithm's precision and robustness based on real scenarios. We test our method using Sentinel-1 and WorldView-3 images, to fit a SAR physical sensor model or correct an existing RPC model by composing it with a complementary transformation.

1.1 Mathematical formulation of the RPC model

The RPC model defines the projection function $\mathcal{P}$ as

$$r_n = \frac{a\left(X_n, Y_n, Z_n\right)}{b\left(X_n, Y_n, Z_n\right)} \quad c_n = \frac{e\left(X_n, Y_n, Z_n\right)}{f\left(X_n, Y_n, Z_n\right)}, \tag{1}$$

[1] Also referred to as Rational Polynomial Coefficients camera model or Rational Function Model in the literature.

where a, b, e, f are cubic polynomials. X, Y, Z represent the longitude, latitude and height of a 3D point; and r, c are the row and column of its projection on the image plane.

Equation 1 uses normalized coordinates for better numerical stability, hence the subscript n. Normalized values are in the range [-1, 1] and they are obtained using two scalars, an offset and a scale factor: $X_n = (X - X_{\text{offset}})/X_{\text{scale}}$, where X could be r, c, X, Y or Z from Equation 1.

Each RPC polynomial p is defined by 20 coefficients as

$$\begin{aligned} p\left(X, Y, Z\right) = \; & p_0 + p_1 Z + p_2 Y + p_3 X + p_4 ZY + p_5 ZX \\ & + p_6 YX + p_7 X^2 + p_8 Y^2 + p_9 Z^2 + p_{10} ZYX \\ & + p_{11} Z^2 Y + p_{12} Z^2 X + p_{13} Y^2 Z + p_{14} Y^2 X \\ & + p_{15} ZX^2 + p_{16} YX^2 + p_{17} Z^3 + p_{18} Y^3 + p_{19} X^3, \end{aligned} \tag{2}$$

where p_i is the i-th coefficient of p. Since we set $p_0 = 1$ for the RPC denominator polynomials, a total of 78 coefficients need to be determined to define a, b, e and f in Equation 1.

2 Related work

RPCs have been used for high-resolution optical satellite imaging since the launch of Ikonos in 1999 [1, 4–6]. In the last decade, they have been proven to be extremely accurate for SAR acquisition systems as well [7, 8].

The RPC model of a satellite image can be constructed using a set of correspondences between image and object space coordinates. Depending on the nature of these correspondences, the literature can be classified into terrain-dependent or independent methods (or a combination of both). Terrain-dependent strategies use Ground Control Points (GCPs), whose object and image coordinates are known in advance relying on manual labeling or on-site measurements. Oppositely, terrain-independent methods derive virtual sets of 2D-3D point correspondences from other geolocation functions, usually a physical sensor model. Once the point correspondences are available, least squares algorithms are typically used to estimate the RPC coefficients that min-

imize the error between the projected 3D points and their image locations.

Several works have underlined the importance of using uniformly distributed points in sufficient amount, covering the different parts of the image and the whole altitude range of the scene [4, 6]. As a result, regularized least squares methods have become widely used to gain robustness to different configurations and enforce well-conditioned normal equations [4, 6, 8, 9]. Additionally, terrain-dependent strategies have explored the selection of optimal and balanced subsets of GCPs, e.g. [4] propose a bucketing strategy or [9] study the benefits of encouraging correspondences located at building edges in urban scenarios. In contrast, terrain-independent methods can arbitrarily generate regular sets of points, but require special care to the boundaries and density of the structure, e.g. [8] investigate the impact of different number of elevation layers for flat and mountainous areas.

3 Method

We follow a terrain-independent approach similar to [4] to fit an RPC model to another input geolocation model. The data used to fit the model consists of a 3D grid of uniformly distributed Control Points (CNPs) within some longitude, latitude and altitude boundaries (Fig. 1). The 2D image point of each CNP can be obtained by projecting it with the input geolocation model, so that each sample results in 5 normalized values, i.e. $(X_i, Y_i, Z_i, r_i, c_i)$. For simplicity, we drop the subscript n of normalized coordinates and replace it by i to refer to the sample index.

Using N CNPs, Equation 1 can be rewritten as a system of equations, in matrix form, following the derivation of [4]:

$$WTI - WG = 0, \tag{3}$$

where

$$\begin{aligned} W &= \mathrm{diag}\left[\frac{1}{b(\boldsymbol{X}_1)}, ..., \frac{1}{b(\boldsymbol{X}_N)}, \frac{1}{f(\boldsymbol{X}_1)}, ..., \frac{1}{f(\boldsymbol{X}_N)}\right] \\ b(\boldsymbol{X}_i) &= b(X_i, Y_i, Z_i), \\ T &= \text{block diag}\left[M_r, M_c\right] \quad M_r, M_c \in \mathbf{R}^{N\times 39} \\ M_{r_i} &= \left[1, Z_i, Y_i, ..., X_i^3, -r_iZ_i, -r_iY_i, ..., -r_iX_i^3\right] \\ M_{c_i} &= \left[1, Z_i, Y_i, ..., X_i^3, -c_iZ_i, -c_iY_i, ..., -c_iX_i^3\right], \\ I &= \left[a_0, ..., a_{19}, b_1, ..., b_{19}, e_0, ..., e_{19}, f_1, ..., f_{19}\right]^T, \\ G &= \left[r_0, ..., r_N, c_0, ..., c_N\right]^T. \end{aligned}$$

In Equation 3, W is a weight matrix with shape $2N \times 2N$, $b(\boldsymbol{X}_i)$ denotes the RPC polynomial b evaluated with the 3D coordinates of the i-th CNP; T is the design matrix with shape $2N \times 78$; I is the solution vector with the 78 RPC coefficients necessary to determine a, b, e and f in Equation 1; and G is a vector of length $2N$ containing the CNPs image coordinates.

Equation 3 can be solved by least squares minimization to estimate I, using the normal equation

$$T^TW^2TI - T^TW^2G = 0. \tag{4}$$

To increase numerical stability, ridge estimation regularization [8, 9] is often added so that the normal equation becomes

$$(T^TW^2T + h^2E)I - T^TW^2G = 0, \tag{5}$$

where E is the identity matrix and h is a scalar controlling the regularization that is applied. To choose the best regularization factor h, the L-curve criterion was introduced in [8]. This heuristic computes the log norm of the solution ($\log \|I\|_h$) versus the log norm of the residual ($\log \|WTI - WG\|_h$) across different values of h that extend from the minimal to the maximal singular value of T. This curve usually has a L-shape, in which the optimum corresponds to the maximum regularization parameter that achieves a small residual. The value corresponding to the corner of the curve, at the position of maximal curvature, is taken to set h automatically. For non-weighted regularized least squares (i.e. weights are set to the identity, $W = E$), the L-curve criterion is fast since the curvature can be computed with closed form expressions [10].

Therefore, we first set $W^{(0)} = E$ (where the superscript denotes the iteration number) and use the L-curve criterion to determine the optimal h and an initial solution $I^{(0)}$. Then, for $i \geq 1$, $W^{(i)}$ is determined

from $I^{(i-1)}$ and is plugged in Equation 5 to solve for $I^{(i)}$ iteratively (the SVD least squares solver is used for stability). The iterations stop when the change in terms of RMSE between the RPC projected CNPs and their image coordinates becomes lower than a tolerance value. After convergence some final ICCV (Iteration by Correcting Characteristic Value [8]) iterations are computed to remove possible biases introduced by the regularization. The same stopping criterion based on the RMSE improvement is used. Each ICCV iteration k can be expressed as

$$(T^T(W^{(k)})^2T+E)I^{(k)} = T^T(W^{(k)})^2G+I^{(k-1)}. \quad (6)$$

4 Experiments

4.1 Data and use cases description

- **SAR.** 32 Sentinel-1 SAR images (Table 1). The data is in Interferometric Wide Swath (IW) mode, each product contains 3 subswaths, and each subswath contains multiple bursts that need to be stitched together to get a continuous image. We construct the Range-Doppler physical sensor model and use the method from Section 3 to fit an equivalent RPC model for each image of the dataset.

- **Optical.** 47 WorldView-3 panchromatic images (Table 1), from the 2016 *IARPA Multi-View Stereo 3D Mapping Challenge* [11]. The original RPCs of the images exhibit small inaccuracies, mainly due to inexact knowledge of the satellite attitude angles, which cause a 3D point to be projected to non corresponding pixels across different images. Bundle adjustment (BA) algorithms are a well-known approach to correct RPC errors [5, 12, 13]. We apply a BA similar to [12] to correc the projection function $\mathcal{P}$ of each RPC into a new $\mathcal{P}_{\text{BA}}$, expressed as

$$\mathcal{P}_{\text{BA}}(X) = \mathcal{P}(R(X-T-C)+C), \quad (7)$$

where X is a 3D point. That is, each RPC is corrected by applying a translation T followed by a rotation R around an approximate camera center C, before applying the original projection $\mathcal{P}$. C is derived by regressing a projective model from each RPC model. Our method from Section 3 is used to fit $\mathcal{P}_{\text{BA}}$ from the composition of $\mathcal{P}$ with T, R, C.

	SAR dataset	**Optical dataset**
platform	S1 A/B	WorldView-3
number of images	32	47
geographic area (lon, lat) center	Albania (18.82,41.02)	Argentina (-58.61,-34.47)
first acquisition date	2019-08-03	2014-11-15
last acquisition date	2020-02-05	2016-01-13
altitude range (m)	$[-533, 2969]$	$[-513, 548]$

Table 1: SAR and optical data used in the experiments. The altitude range of the area covered by each collection of images is defined using the [min, max] values from the corresponding SRTM digital elevation model ±500 m to consider tall buildings or fine irregularities beyond bare ground level.

4.2 Performance assessment

To apply the method described in Section 3, the altitude limits of the grid of CNPs are set using the altitude ranges in Table 1. To assess the RPC fitting, we use a grid of Check Points (CKPs), which are located in the middle of each pair of consecutive CNPs. The RMSE between the image coordinates obtained by projecting the CKPs with the output RPC and the image coordinates obtained using the input geolocation model, measured in pixels, is used as evaluation metric. We set a tolerance of 10^{-10} for the stopping criterion based on the RMSE improvement, and a maximum of 20 iterations for the weighted least squares and the ICCV iterations.

Two type of experiments were conducted to assess the performance and robustness of the method:

- *Varying surface area.* For each image, we fit different RPCs by gradually increasing the longitude and latitude limits of the grid of CNPs from a small square centered at the image center to a larger square including the entire image. The number of CNPs is fixed, with 50×50 samples in the longitude and latitude dimensions and 10 elevation layers (25000 points in total).

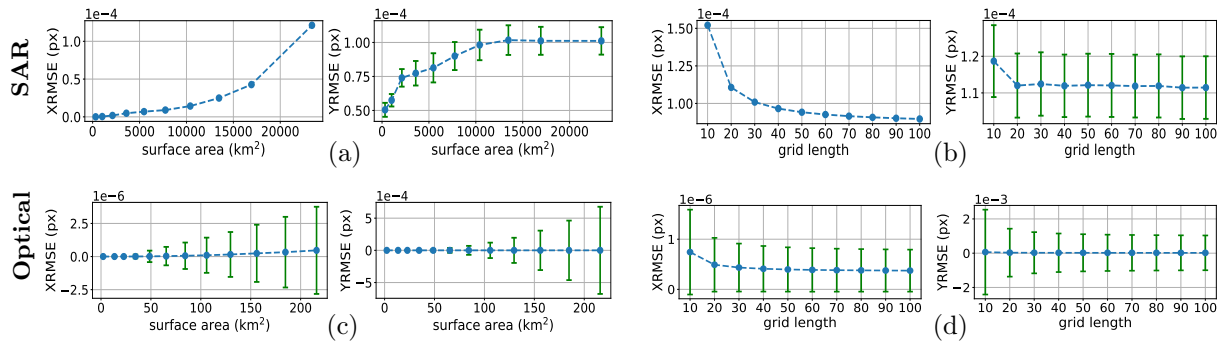

Figure 2: RPC fitting error varying grid length and surface area on the SAR (a-b) and optical data (c-d) described in Section 4.1. Each vertical bar corresponds to the $[-\sigma/2,\ \mu,\ \sigma/2]$ values of the root-mean-square error (RMSE) evaluated across the different images of the datasets, in each dimension of the image plane, where μ corresponds to the mean and σ to the standard deviation.

- *Varying grid length.* For each image, we fit different RPCs, by increasing the number of CNPs in each elevation layer, i.e. $n \times n$ where n is the *grid length*. The longitude and latitude boundaries of the grid are fixed using the equivalent limits of the image plane.

Overall, the results presented in Fig. 2 show that the RPCs constructed with our method approximate the geolocation models with very high accuracy for the two datasets and the two scenarios outlined in Section 4.1. The RMSE is in the order of 10^{-4} pixels or less in both dimensions of the image plane for the majority of configurations that were tested, which emphasizes the robustness of the method.

The experiments with different grid lengths (Fig. 2b and 2d) show that 10 samples in the longitude and latitude directions is already a good choice and increasing this number beyond 20 does not result in significant improvements. The experiments with varying surface area (Fig. 2a and 2c) show that both the overall RMSE values and its variation across the different images increase with the size of the area being fitted. This is probably due to the fact that the SAR physical sensor model is less smooth for large neighborhoods. A similar behavior is obtained with the optical dataset, where the original RPCs are known to behave locally as an affine camera [5, 13].

5 Conclusion

This article described an automatic algorithm to fit the RPC model of a satellite image in a terrain-independent manner. The inputs of the method are a regular grid of 3D points (CNPs), with multiple elevation layers, and the 2D locations of the points on the image plane. We evaluated the method on real scenarios using collections of SAR and optical satellite images, and assessed its performance by varying the CNPs configuration. Finally, we release an open-source implementation of the algorithm as an easy-to-use Python package.

Acknowledgements

Work partly financed by IDEX Paris-Saclay IDI 2016, ANR-11-IDEX-0003-02, Office of Naval research grant N00014-17-1-2552 and N00014-20-S-B001, DGA Astrid project « filmer la Terre » n° ANR-17-ASTR-0013-01, MENRT, and by a grant from Région Île-de-France.

References

[1] J. Grodecki, "IKONOS stereo feature extraction - RPC approach," in *ASPRS Annual Confer-*

ence, 2001.

[2] C. de Franchis, E. Meinhardt-Llopis, D. Greslou, and G. Facciolo, "Attitude refinement for orbiting pushbroom cameras: a simple polynomial fitting method," *Image Processing On Line*, vol. 2015, pp. 328–361, 2015.

[3] J.C. Curlander, "Location of spaceborne SAR imagery," *TGRS*, vol. GE-20, no. 3, pp. 359–364, 1982.

[4] C.V. Tao and Y. Hu, "A comprehensive study of the rational function model for photogrammetric processing," *Photogrammetric Engineering & Remote Sensing*, vol. 67-12, 2001.

[5] C.S. Fraser, G. Dial, and J. Grodecki, "Sensor orientation via RPCs," *ISPRS*, vol. 60, no. 3, pp. 182–194, 2006.

[6] T. Long, W. Jiao, and G. He, "RPC estimation via ℓ_1-norm regularized least squares (L1LS)," *TGRS*, vol. 53, no. 8, 2015.

[7] G. Zhang, W. Fei, Z. Li, X. Zhu, and D. Li, "Evaluation of the RPC model for spaceborne SAR imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 76-6, pp. 727–733, 2010.

[8] L. Zhang, X. He, T. Balz, X. Wei, and M. Liao, "Rational function modeling for spaceborne SAR datasets," *ISPRS*, vol. 66, no. 1, pp. 133–145, 2011.

[9] Y. Wang, Y. Zhang, and N. Su, "RPC estimation via feature points for urban areas," in *IGARSS*, 2016, pp. 6684–6687.

[10] P.C. Hansen and D.P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.

[11] M. Bosch, Z. Kurtz, S. Hagstrom, and M. Brown, "A multiple view stereo benchmark for satellite imagery," in *AIPR Workshop*, 2016, pp. 1–9.

[12] R.A. Beyer, O. Alexandrov, and S. McMichael, "The Ames Stereo Pipeline: NASA's open source software for deriving and processing terrain data," *Earth and Space Science*, vol. 5, no. 9, pp. 537–548, 2018.

[13] R. Marí, C. de Franchis, E. Meinhardt-Llopis, and G. Facciolo, "To bundle adjust or not: A comparison of relative geolocation correction strategies for satellite multi-view stereo," in *ICCV Workshops*, 2019.