*Article*

# Improvement of EPIC/DSCOVR Image Registration by Means of Automatic Coastline Detection

**Víctor Molina García \***[ID]**, Sruthy Sasi**[ID]**, Dmitry S. Efremenko**[ID] **and Diego Loyola**[ID]

Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany
\* Correspondence: Victor.MolinaGarcia@dlr.de

**Abstract:** In this work, we address the image geolocation issue that is present in the imagery of EPIC/DSCOVR (Earth Polychromatic Imaging Camera/Deep Space Climate Observatory) Level 1B version 2. To solve it, we develop an algorithm that automatically computes a registration correction consisting of a motion (translation plus rotation) and a radial distortion. The correction parameters are retrieved for every image by means of a regularised non-linear optimisation process, in which the spatial distances between the theoretical and actual locations of chosen features are minimised. The actual features are found along the coastlines automatically by using computer vision techniques. The retrieved correction parameters show a behaviour that is related to the period of DSCOVR orbiting around the Lagrangian point $L_1$. With this procedure, the EPIC coastlines are collocated with an accuracy of about 1.5 pixels, thus significantly improving the original registration of about 5 pixels from the imagery of EPIC/DSCOVR Level 1B version 2.

**Keywords:** EPIC; registration; geolocation; computer vision; regularisation

## 1. Introduction

The spacecraft DSCOVR (Deep Space Climate Observatory) provides a unique view of the Earth from its Lissajous orbit around the Earth–Sun Lagrangian point $L_1$, at a distance of 1.5 million kilometres from the Earth. DSCOVR carries a range of sensors on board, including EPIC (Earth Polychromatic Imaging Camera) for climate science applications [1].

EPIC consists of a CCD (Charge-Coupled Device) camera that monitors the sunlit half of the Earth in 10 spectral channels ranging in the ultraviolet, visible and near infrared—approximately from 317 to 780 nm. Every two hours, EPIC measures its 10 channel images with a shape of $2048 \times 2048$ detector pixels, and a varying ground pixel size of approximately $12 \times 12$ km$^2$ at the center of the images. These available channels make EPIC a suitable candidate for several applications, such as monitoring the vegetation condition [2], the synoptic ozone [3], the sulphur dioxide content from volcanic eruptions [4,5], the aerosol layer height and optical depth [6,7], and different cloud properties [8]. Extensive work has been performed to estimate the cloud information content in the EPIC oxygen bands [9,10], as well as its sensitivity to liquid-phase cloud microphysical parameters [11].

Standard techniques for retrieval of atmospheric properties, such as the layer height and the optical depth of clouds and aerosols, are non-linear least squares fitting [12] and Tikhonov regularisation [13], which match the results of forward simulations (i.e., synthetic spectra) with the actual measurements (i.e., real spectra). The forward simulations are performed on the base of radiative transfer models. Due to the particular viewing geometry of EPIC, with scattering angles ranging between 168 and 176 degrees, the radiative transfer modelling faces challenges in terms of computational efficiency and accuracy [14,15], which can be solved by using various acceleration techniques [16–18]. Furthermore, this approach needs to assume that the surface properties required as input for the radiative transfer model, e.g., the surface altitude and the spectral surface albedo,

are already known from external sources. The assignment of surface altitude and surface albedos must be performed for every image pixel based on their corresponding latitude and longitude values allocated in the EPIC Level 1B (L1B) arrays of Earth coordinates, which are provided together with the measurement arrays.

Since the evaluation of these surface properties for each image pixel is based on its geolocation information, any incorrect registration of the images can lead into an invalid assignment of these surface properties and, thus, an erroneous retrieval of the atmospheric properties. In the case of EPIC, the geolocation arrays, including the latitude and longitude locations as well as the Sun and instrument viewing angles, are computed by means of a complete navigation algorithm. Essentially, it determines the spacecraft location and orientation and maps the 3D-model coordinates into the 2D image coordinates [19]. This mapping includes an optical correction modelled as a small barrel distortion. Although this navigation algorithm provides a good first estimate for the image geolocation, the imagery from EPIC L1B version 2 presents a misregistration that is especially noticeable when inspecting the expected and actual locations of the land bodies (Figure 1). A similar problem was already mentioned and briefly described for the previous release L1B version 1 [20].

In this article, we develop an automatic registration algorithm for EPIC that relies on the expected location of the Earth land bodies. Assuming that the geolocation error is smaller than 10 detector pixels, we state that it is enough to apply a transformation on the measurement pixel indices, consisting of a motion (translation plus rotation of the Earth's disk) together with a small optical correction, in order to improve the registration of the EPIC imagery. The optimal transformation parameters must be computed for every EPIC L1B dataset with the help of reference pixels taken from the expected and actual locations of the land bodies, which can be determined by using conventional pattern recognition techniques that are of common use in computer vision.
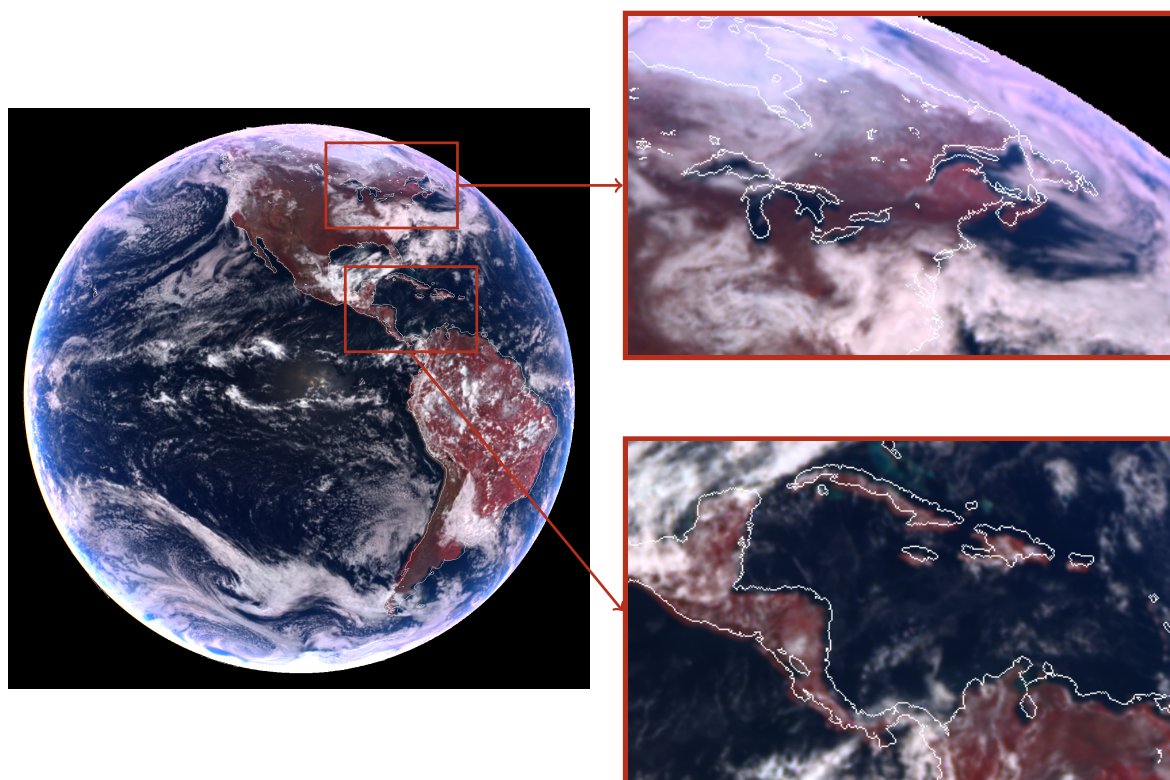


**Figure 1.** Example of reflectance image from EPIC L1B version 2 on 20 March 2016 at 18:36:56 UTC, with the coastline shape vector drawn in white line. On the left side, the full Earth's disk is shown. On the right side, we zoom into two regions where the misregistration is noticeable.

## 2. Methodology

In this section, we present a mathematical model that describes the misregistration of the EPIC L1B images, and we also describe the procedure to estimate the optimal transformation parameters that improve their registration quality. The original EPIC L1B images are accessible through NASA Earth Data portal [21]. In the following, we use the word *image* to refer to false-RGB images containing the reflectance values from the EPIC channels of 779.5, 551.0 and 443.0 nm, with the northern polar region of the Earth's disk located at the top of the image (as the example shown in Figure 1). The original counts per second measured by EPIC can be converted into reflectance by using the calibration factors estimated by Geogdzhayev and Marshak [22] and dividing then by the solar zenith-directional cosine. Although the transformation parameters were obtained for these false-RGB images, they can be used for all the EPIC L1B channels.

### 2.1. Mathematical Description

Given an original image pixel located at indices $[x_d, y_d]^\top$ (hereinafter the superscript $\top$ stands for "transpose") which is not aligned with the proper pixels from the arrays of Earth coordinates, we model its transformation into a registered pixel at indices $[x_r, y_r]^\top$ by the following expression:

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_u - x_0 \\ y_u - y_0 \end{bmatrix}, \tag{1}$$

where $[x_s, y_s]^\top$ is the shift vector, $\theta$ is the rotation angle around the center of rotation $[x_0, y_0]^\top$, and $[x_u, y_u]^\top$ are the pixel indices after radial distortion correction by means of the single-parameter Fitzgibbon division model [23,24]:

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + g(\lambda, r) \begin{bmatrix} x_d - x_c \\ y_d - y_c \end{bmatrix}, \qquad g(\lambda, r) = \frac{1}{1 + \lambda r^2}, \tag{2}$$

where $\lambda$ is the first radial distortion coefficient, $[x_c, y_c]^\top$ is the center of distortion and $r$ is the distance from the distorted point to the center of distortion:

$$r = \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2}. \tag{3}$$

The alignment transformation (Equation (1)) depends on eight different parameters which are unknown and need to be estimated. Before proceeding further, and based on empirical observations, we need to apply two restrictions in order to reduce the complexity of the problem:

1. The center of distortion and the center of rotation are assumed identical, i.e., $x_0 = x_c$ and $y_0 = y_c$.
2. The center of distortion/rotation is assumed to be known and located at the center of the image.

Under these assumptions, the transformation (Equation (1)) from $[x_d, y_d]^\top$ to $[x_r, y_r]^\top$ can be simplified into:

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} g(\lambda, r)\,(x_d - x_c) \\ g(\lambda, r)\,(y_d - y_c) \end{bmatrix}, \tag{4}$$

which depends on four parameters (i.e., $x_s$, $y_s$, $\theta$, and $\lambda$), as the point $[x_c, y_c]^\top$ is fixed now. To compute the optimal transformation parameters, we adapt Equation (4) for $n > 2$ known pairs of distorted and registered points $\{([x_{di}, y_{di}]^\top, [x_{ri}, y_{ri}]^\top)\}_{i=1}^n$, i.e.,

$$\begin{bmatrix} x_{r1} \\ y_{r1} \\ \vdots \\ x_{rn} \\ y_{rn} \end{bmatrix} = \begin{bmatrix} x_s + x_c \\ y_s + y_c \\ \vdots \\ x_s + x_c \\ y_s + y_c \end{bmatrix} + \begin{bmatrix} \cos\theta & \sin\theta & \cdots & 0 & 0 \\ -\sin\theta & \cos\theta & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cos\theta & \sin\theta \\ 0 & 0 & \cdots & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} g(\lambda, r_1)\,(x_{d1} - x_c) \\ g(\lambda, r_1)\,(y_{d1} - y_c) \\ \vdots \\ g(\lambda, r_n)\,(x_{dn} - x_c) \\ g(\lambda, r_n)\,(y_{dn} - y_c) \end{bmatrix}, \tag{5}$$

rewritten in compact form as

$$z_{\mathrm{r}} = f(z_{\mathrm{d}}, z_{\mathrm{c}}, p) \tag{6}$$

where $f$ denotes the transformation of the distorted data vector $z_{\mathrm{d}} = [x_{\mathrm{d}1}, y_{\mathrm{d}1}, \ldots, x_{\mathrm{d}n}, y_{\mathrm{d}n}]^{\mathsf{T}}$ into the registered data vector $z_{\mathrm{r}} = [x_{\mathrm{r}1}, y_{\mathrm{r}1}, \ldots, x_{\mathrm{r}n}, y_{\mathrm{r}n}]^{\mathsf{T}}$ by means of a known center of distortion/rotation $z_{\mathrm{c}} = [x_{\mathrm{c}}, y_{\mathrm{c}}]^{\mathsf{T}}$ and an unknown state vector $p = [x_{\mathrm{s}}, y_{\mathrm{s}}, \theta, \lambda]^{\mathsf{T}}$. The state vector $p_{\alpha}$ that optimally transforms $z_{\mathrm{d}}$ into $z_{\mathrm{r}}$ can be computed by least squares minimisation of the Tikhonov function $\mathcal{T}(p)$ defined as

$$\mathcal{T}(p) = \frac{1}{2}\left( ||f(z_{\mathrm{d}}, z_{\mathrm{c}}, p) - z_{\mathrm{r}}||^2 + \alpha ||L(p - p_{\mathrm{a}})||^2 \right), \tag{7}$$

where $\alpha$ is the regularisation parameter, $L$ is the regularisation matrix, and $p_{\mathrm{a}}$ is the a priori state vector. The procedure to minimise $\mathcal{T}(p)$ is iterative; starting from an initial guess $p_{(0)}$, the $k$th iteration consists of four steps [25]:

1.　Compute the vector $y_{(k)}$ at iteration $k$ as

$$y_{(k)} = z_{\mathrm{r}} - f(z_{\mathrm{d}}, z_{\mathrm{c}}, p_{(k)}) + J_{(k)}(p_{(k)} - p_{\mathrm{a}}), \tag{8}$$

where $J_{(k)}$ denotes the Jacobian matrix of $f$ with respect to the state vector $p$ at iteration $k$:

$$J_{(k)} = \left[\frac{\partial f_i}{\partial p_j}\right]_{p = p_{(k)}}, \qquad i = 1, \ldots, 2n, \quad j = 1, \ldots, 4. \tag{9}$$

2.　Compute the regularised generalised inverse $J^{\dagger}$ at iteration $k$ by means of generalised singular value decomposition:

$$J^{\dagger}_{(k)} = (J^{\mathsf{T}}_{(k)} J_{(k)} + \alpha L^{\mathsf{T}} L)^{-1} J^{\mathsf{T}}_{(k)}, \tag{10}$$

3.　Compute the state vector $p$ for the next iteration $k + 1$:

$$p_{(k+1)} = p_{\mathrm{a}} + J^{\dagger}_{(k)} y_{(k)}. \tag{11}$$

4.　Check the convergence criteria. If any is passed, set $p_{\alpha} = p_{(k+1)}$ and exit; otherwise, go to Step 1 for iteration $k + 1$. We use the following convergence criteria:

  - X-convergence criterion: $||p_{(k+1)} - p_{(k)}|| < \delta$, where $\delta$ is a predefined tolerance value.
  - S-convergence criterion: $|s_{(k+1)} - s_{(k)}| < \varepsilon$, where $\varepsilon$ is a predefined tolerance value and $s_{(k)} = ||f(z_{\mathrm{d}}, z_{\mathrm{c}}, p_{(k)}) - z_{\mathrm{r}}||^2$ is the squared residual sum at iteration $k$.

Our fitting problem is much more sensitive to the parameters $[\theta, \lambda]^{\mathsf{T}}$, i.e., small perturbations in these parameters cause large transformations, especially for points far from the center of distortion/rotation $[x_{\mathrm{c}}, y_{\mathrm{c}}]^{\mathsf{T}}$. For this reason, the computation of the optimal state vector $p_{\alpha}$ is performed in two steps:

1.　Define an a priori state vector $p_{\mathrm{a}0} = [0, 0, \theta_{\mathrm{a}}, \lambda_{\mathrm{a}}]^{\mathsf{T}}$ with no shift, and compute the optimal state vector $p_{\alpha 0}$ by applying Tikhonov least squares minimisation with $\theta = \theta_{\mathrm{a}}$ and $\lambda = \lambda_{\mathrm{a}}$ as fixed parameters.

2.　Use $p_{\alpha 0}$ from the previous step as the new a priori state vector $p_{\mathrm{a}} = p_{\alpha 0}$, and compute the optimal state vector $p_{\alpha}$ by applying Tikhonov least squares minimisation with no fixed parameters.

The configuration for the second-step Tikhonov least squares minimisation is summarised as follows:

1.  The analysis of the Jacobian matrices at the first iteration by means of the generalised singular value decomposition shows that $\alpha \approx 100$ is an appropriate value for the regularisation parameter.
2.  The regularisation matrix $L$ is defined as a diagonal matrix,

$$L = \mathrm{diag} \left[ \frac{w_{x_s}}{\varepsilon_{x_s}}, \frac{w_{y_s}}{\varepsilon_{y_s}}, \frac{w_\theta}{\varepsilon_\theta}, \frac{w_\lambda}{\varepsilon_\lambda} \right],$$

where, for every parameter, we denote by $\varepsilon$ its expected dispersion, and we denote by $w \in [0, +\infty)$ a weighting factor that determines its freedom in the retrieval, so that $w \to 0^+$ implies total freedom and $w \to +\infty$ fixes the parameter to the a priori solution [25]. Based on the expected dispersions of the state vector components, which are set as

$$\varepsilon_{x_s} = 10, \qquad \varepsilon_{y_s} = 10, \qquad \varepsilon_\theta = 0.1 \text{ deg}, \qquad \varepsilon_\lambda = 1 \times 10^{-8},$$

the optimal weights have to be found empirically by performing a set of retrievals with synthetic data, ensuring that the state vector components stay within the valid ranges around the a priori solution. For our case, the following optimal values are found:

$$w_{x_s} = 0, \qquad w_{y_s} = 0, \qquad w_\theta = 10, \qquad w_\lambda = 10,$$

thereby implying total freedom for $[x_s, y_s]^\mathsf{T}$ and reduced freedom for $[\theta, \lambda]^\mathsf{T}$.

3.  The first a priori state vector $p_{a0}$ is set with the following initial guesses for the rotation angle and the radial distortion parameter:

$$\theta_a = 0.5 \text{ deg}, \qquad \lambda_a = -5 \times 10^{-9}.$$

The workflow for the complete non-linear optimisation process is summarised in Figure 2a.

*2.2. Detection of Matching Coastline Features*

The registration of an image by means of the model described in Section 2.1 assumes that the data vectors $z_d$ and $z_r$ are known. In our case, they can be built with a sufficient number of pairs of incorrect points $[x_d, y_d]^\mathsf{T}$ and correct points $[x_r, y_r]^\mathsf{T}$ within the image. The coastline that is visible in this kind of images is a source of these points of interest, as the correct coastline can be determined with the arrays of Earth coordinates, which are available. With this consideration, we compute the required pairs of incorrect and correct points as follows (see Figure 2b):

1.  Create a mask with the theoretical coastline, i.e., that inferred from the arrays of Earth coordinates.
2.  Create a mask with the radiometric coastline, i.e., that inferred from the actual image.
3.  Find pairs of common features from both coastline masks by using computer vision techniques.

2.2.1. Computation of the Theoretical Coastline

The computation of the theoretical coastline, i.e., the coastline that is assumed to be correct, can be divided into three steps:

1.  Create the theoretical land mask by checking for every image pixel if its (latitude, longitude) pair is contained inside a land polygon from the low-resolution GSHHG (Global Self-consistent, Hierarchical, High-resolution Geography database) [26].
2.  Apply one morphological binary erosion to the theoretical land mask with a flat diamond shaped structuring element of dimensions $3 \times 3$ [27].
3.  Compute the theoretical coastline mask as the result of the bitwise operator XOR on the original and eroded theoretical land masks.
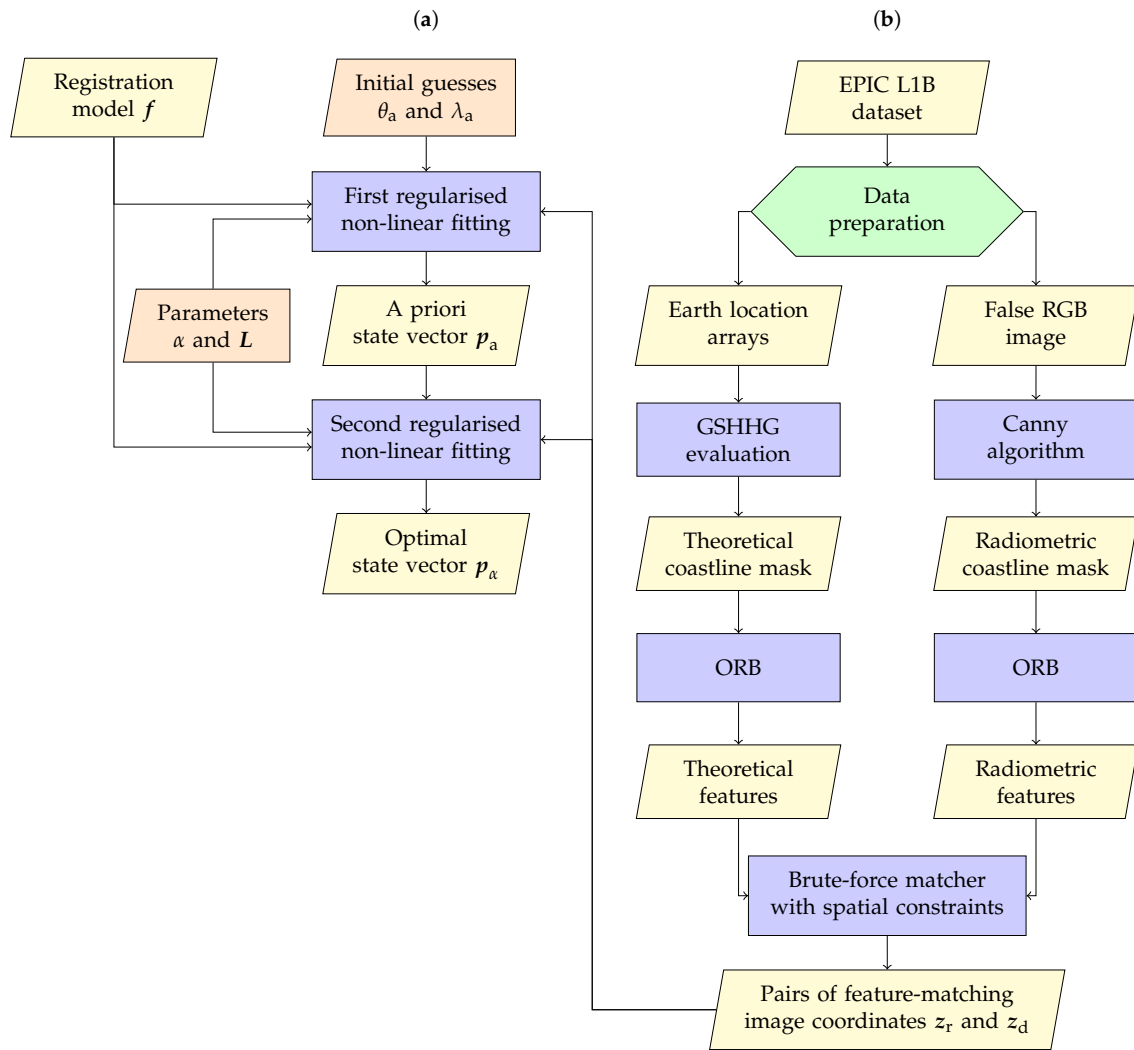
**Figure 2.** Workflow chart for the computation of the optimal transformation parameters for a given image. (**a**) Main workflow chain in which the non-linear least squares fitting with regularisation is performed. (**b**) Secondary workflow chain in which the known data vectors $z_r$ and $z_d$ are prepared.

### 2.2.2. Computation of the Radiometric Coastline

The computation of the radiometric coastline, i.e., the coastline that is visible in the image, is a common edge detection problem in image analysis, and can be performed as follows:

1.  Convert the image channel with the highest contrast between land and water into 8-bit form (i.e., the false-red channel, which corresponds to the 779.5 nm EPIC channel).
2.  Compute the median value $v$ from the Earth pixels of this 8-bit image.
3.  Compute the radiometric coastline mask by applying the Canny edge detection algorithm [28] on the previous 8-bit image with hysteresis thresholding parameters given by

$$t_{\text{lower}} = \max[0, (1-\sigma)v], \qquad t_{\text{upper}} = \min[255, (1+\sigma)v], \tag{12}$$

where $\sigma \geq 0$ is an argument that controls the separation between both thresholds [29] and its value is found empirically: if $\sigma$ is small, only the strong borders are preserved; if $\sigma$ is big, both strong and weak borders are kept. For our purpose, we found that $\sigma = 0.33$ is a good compromise.

With this procedure, the radiometric coastline will be polluted in general with other edges also present in the image, e.g., cloud edges (see Figure 3). However, they do not need to be filtered as long as the detected matching features are restricted to the neighbourhood of the theoretical coastline.
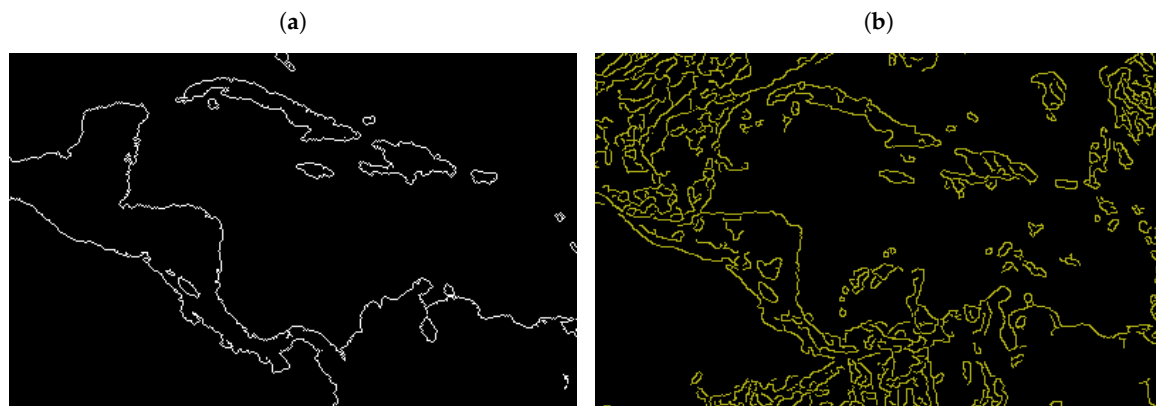
(**a**)　　　　　　　　　　　　　　　　　　　　　(**b**)



**Figure 3.** Comparison of the (**a**) theoretical and (**b**) radiometric coastlines for one of the regions shown in Figure 1. The radiometric coastline is partially polluted with other type of borders detected by the Canny algorithm (e.g., clouds or interfaces between two different land covers).

### 2.2.3. Matching of Coastline Features

Once the theoretical and radiometric coastline masks are computed, it is possible to find features that are common to both masks, as long as the misalignment and distortion of the radiometric coastline is not severe. Such image matching problem is one of the fundamental research topics in computer vision [30]; the procedure to find common features between two images consists of three steps:

1. Detect keypoints (e.g., edges, corners, and regions of interest) in the compared images.
2. Describe every keypoint by a descriptor vector with information from its neighbourhood.
3. Match keypoint pairs based on the similarity of their descriptor vectors.

Examples of successful detectors and descriptors are SIFT (Scale-Invariant Feature Transform) [31], SURF (Speeded-Up Robust Features) [32] and ORB (Oriented FAST and Rotated BRIEF) [33]. OpenCV [34] already provides a complete interface to these algorithms. For our purpose, we use ORB because it is not patented, its time performance is better, and the amount of detected keypoints is enough. ORB consists of two routines:

1. The ORB detector is a modified version of the keypoint detector FAST (Features from Accelerated Segment Test) [35]. In addition to the original FAST, it also computes the orientation angle of the detected keypoints.
2. The ORB descriptor is a modified version of the descriptor BRIEF (Binary Robust Independent Elementary Features) [36]. The binary descriptor vector generated by the original BRIEF shows problems when identifying matching keypoints under rotation conditions; ORB fixes the issue taking into consideration the orientation computed by the ORB detector.

The keypoints detected and described with ORB for both the theoretical and radiometric coastline masks are matched by using a brute-force matcher with Hamming distance as measurement (see [37] for further details). Given a set of theoretical keypoints $\{z_{1k_1}\}_{k_1=1}^{n_1}$ with descriptor vectors $\{d_{1k_1}\}_{k_1=1}^{n_1}$ and a set of radiometric keypoints $\{z_{2k_2}\}_{k_2=1}^{n_2}$ with descriptor vectors $\{d_{2k_2}\}_{k_2=1}^{n_2}$, the matching algorithm proceeds for every theoretical keypoint $z_{1i}$, $1 \leq i \leq n_1$, as follows:

1.  Compute the Hamming distance (i.e., the number of positions at which the corresponding values of arrays are different [38]) $h_{ik_2}$ from $z_{1i}$ to every radiometric keypoint $z_{2k_2}$, $1 \leq k_2 \leq n_2$, as

$$h_{ik_2} = \sum_{m=1}^{n_{\text{bits}}} (\boldsymbol{d}_{1i})_m \oplus (\boldsymbol{d}_{2k_2})_m, \tag{13}$$

where $n_{\text{bits}}$ is the length of the binary descriptor vectors, $\oplus$ is the XOR operator, and $(\boldsymbol{d}_{1i})_m$ and $(\boldsymbol{d}_{2k_2})_m$ denote the $m$th components of the descriptor vectors $\boldsymbol{d}_{1i}$ and $\boldsymbol{d}_{2k_2}$, respectively.

2.  Select the radiometric keypoint $z_{2j}$ with minimum Hamming distance to $z_{1i}$ as the matching candidate for the theoretical keypoint $z_{1i}$, where

$$j = \arg\min\{h_{ik_2}\}_{k_2=1}^{n_2}. \tag{14}$$

3.  For the given radiometric keypoint $z_{2j}$, compute its Hamming distance $h_{k_1j}$ to every theoretical keypoint $z_{1k_1}$, $1 \leq k_1 \leq n_1$, as

$$h_{k_1j} = \sum_{m=1}^{n_{\text{bits}}} (\boldsymbol{d}_{1k_1})_m \oplus (\boldsymbol{d}_{2j})_m. \tag{15}$$

4.  The pair $(z_{1i}, z_{2j})$ is a valid matching pair only if $z_{1i}$ is the theoretical keypoint with minimum Hamming distance to $z_{2j}$, otherwise the pair is discarded, i.e., it is valid only if

$$i = \arg\min\{h_{k_1j}\}_{k_1=1}^{n_1}. \tag{16}$$

The previous procedure ensures that all the valid pairs of matching keypoints hold the minimum Hamming distance criterion reciprocally. Note that this criterion, however, does not consider any spatial restriction, but only the similarity between the descriptor vectors. Due to the spatial constraints of our problem, the valid pairs need to be filtered based on two additional spatial criteria (Figure 4):

1.  The spatial distance between a theoretical keypoint $z_{1i}$ and the theoretical coastline.
2.  The spatial distance between the keypoints of a matching pair $(z_{1i}, z_{2j})$.
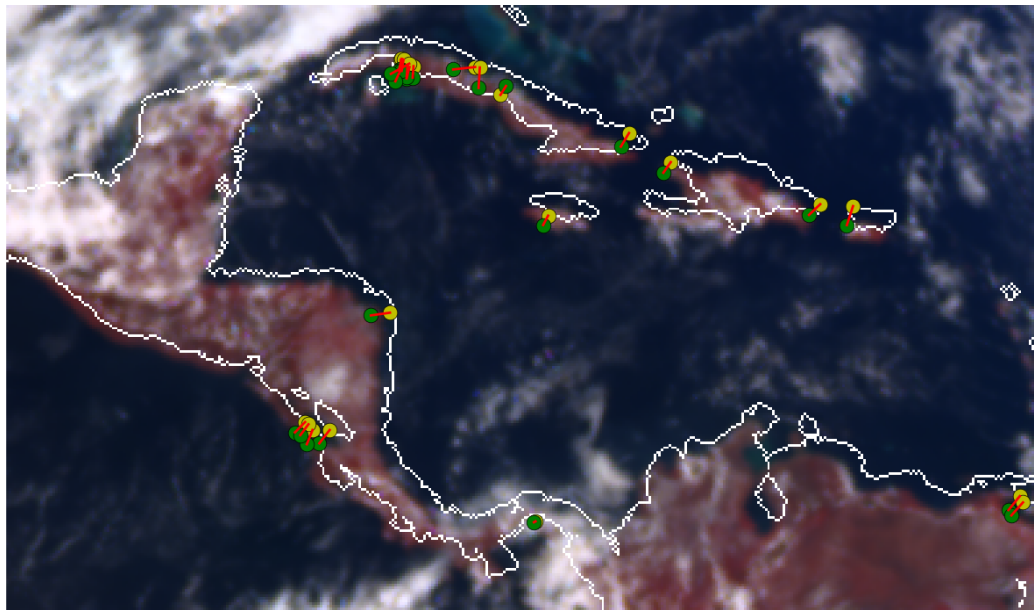


**Figure 4.** Pairs of matching keypoints found within one of the regions shown in Figure 1 after applying the outlier rejection step based on spatial criteria. The keypoints detected in the theoretical coastline (white lines) are shown as yellow points, while the corresponding keypoints detected in the radiometric coastline are shown as green points. Every matching pair is connected with a red segment.

If any of these spatial distances is greater than a threshold (e.g., 10 pixels), the pair is discarded. After the application of this filter, the remaining pairs are used to build the registered data vector $z_r$ with the theoretical keypoints and the distorted data vector $z_d$ with the radiometric keypoints. Alternatively, an approach based on RANSAC (RANdom SAmple Consensus) [39] could be added afterwards in the retrieval procedure to generate $z_r$ and $z_d$. We do not use it in this article because the improvement in our registration problem was small and it increased the computation time in excess.

## 3. Results

We performed the estimation of the optimal transformation parameters for all the datasets from EPIC L1B version 2 within the time period from 13 June 2015 to 31 July 2018. The parameters were computed with an Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz processor using a single thread, and it required approximately 15 s of computation time for every image (i.e., roughly between 18 and 36 h for every year of datasets, assuming 12–24 available datasets per day). To assess the behaviour of the obtained results, we analysed the following aspects:

1.  The performance of the non-linearised fitting procedure in reducing the spatial distance between the image coordinates from matching theoretical and radiometric features.
2.  The global impact of this correction procedure on the image registration quality.
3.  The behaviour of the retrieved transformation parameters as a function of time.

In Figure 5, we show the result of computing the spatial distances between the points of every matching pair before and after applying the registration procedure described in the previous section. The histograms were calculated for all the matching pairs that were found in all the EPIC L1B images within the complete time period under study. In Figure 5a, we can observe a normal distribution of spatial distances spread within the interval of $[0, 10]$ pixels, with a higher density of occurrences in the subinterval of $[3, 7]$ pixels. We can find 50% of the values in the range of $[0, 5]$ pixels. The histogram is cut for values higher than 10 pixels due to the spatial constraints defined in Section 2.2.3. In Figure 5b, we see that the distribution of spatial distances is turned into a Gamma distribution whose mode is found in the bin of $[1.25, 1.50]$ pixels, and where approximately 50% of the values are located within the range of $[0, 1.75]$ pixels. The distribution also shows a long decreasing tail on its right side, meaning that a small percentage of values are now greater than 5 pixels.
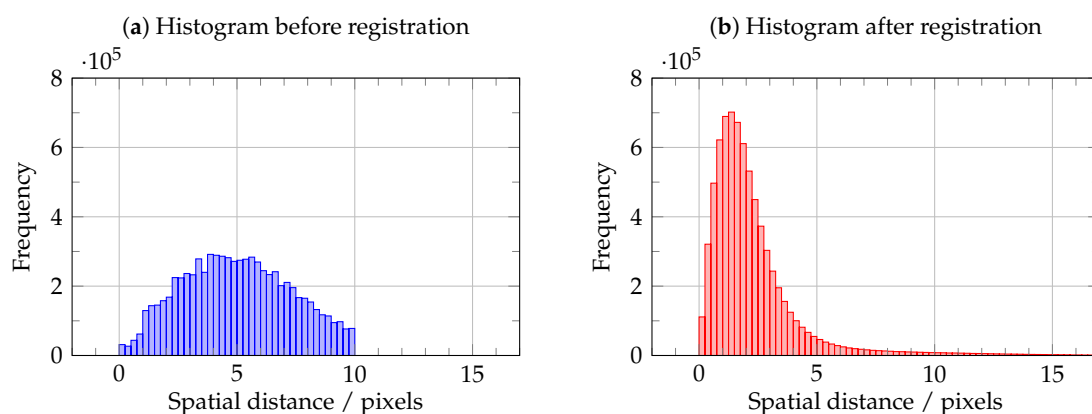


**Figure 5.** Histograms of spatial distances (in pixels) between the points of matching pairs: (**a**) before applying registration, i.e., as they are originally located; and (**b**) after applying registration, i.e., after transforming the distorted points into registered points with the model described in Section 2. The histograms were computed for all the matching pairs found in three years of EPIC L1B images.

From these histograms, we conclude that the transformation significantly reduces the spatial distance between pairs of matching points. As long as these matching pairs are representative of the distortion that the original images suffer, this fact translates into an improvement in the registration

of the images. The existence of the right tail in the histogram in Figure 5b also suggests that some matching pairs do not reduce its spatial distance after the transformation, or they might even increase it. Indeed, this means that there are matching pairs whose points do not correspond to the same feature, but they could still pass the quality filters applied in our procedure. As the amount of wrong matching pairs is not substantial, they do not significantly influence in the retrieval of the transformation parameters, and for every wrong matching pair the points are moved to their new positions determined by the optimal transformation parameters. The location of the histogram mode in Figure 5b is above one pixel due to the necessary discretisation of the coastline together with the actual ground pixel size. The theoretical coastlines were always computed as the last contour of land pixels for every land body based on the GSHHG database. The radiometric coastlines were obtained by finding intensity borders within the false-RGB images, which coincide in general with the actual land borders (see Figure 6), but in some cases they can be found one pixel inside land (because the land–ocean interface pixels are a mixture of land and ocean information) or one pixel outside land (due to water turbidity).
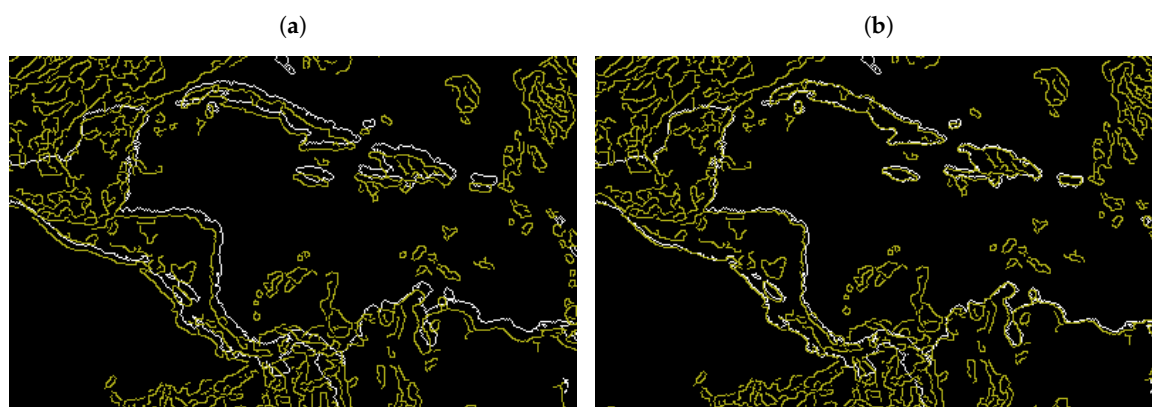
(**a**) (**b**)



**Figure 6.** Comparison of the theoretical coastline (white lines) and the radiometric coastline (yellow lines) for one of the regions shown in Figure 1 (**a**) before and (**b**) after applying the registration.

In Figure 7, we show the application of the registration procedure to four examples of EPIC L1B image regions. On the left side, we observe the original L1B images with the corresponding theoretical coastline shape vector superposed. On the right side, the image pixels are relocated by means of our registration model with the transformation parameters automatically retrieved for every image. It can be observed that there is an improvement in the collocation of the radiometric coastline and the theoretical coastline with respect to their original situation. Because the features were detected along the coastlines, the density of matching features is sensitive to the percentage of land that is visible on every image. As a consequence, the procedure may fail for those images in which the majority of the Earth's disk corresponds to the Pacific Ocean due to the absence of matching features with enough quality. For these situations, we must rely on the retrieved parameters from other images that are close in time.

In Figure 8, we show the time evolution of the transformation parameters for the complete time period under study. The time series was computed for the daily mean values of each parameter. We observed that the horizontal shift $x_s$ has an oscillatory behaviour around its mean value of 2.5 pixels, while the vertical shift $y_s$ also presents a periodic response around its mean value of $-0.2$ pixels. The rotation angle $\theta$ and the distortion parameter $\lambda$ present noisier time series, and their values remain closer to the mean values of 0.498 deg and $-4.958 \times 10^{-9}$, respectively. Indeed, this means that these two parameters tend to stay close to their a priori solutions—an expected result based on the selection of the regularisation matrix $L$. The negative sign of $\lambda$ indicates that the original EPIC L1B images still suffer from a slight barrel distortion.
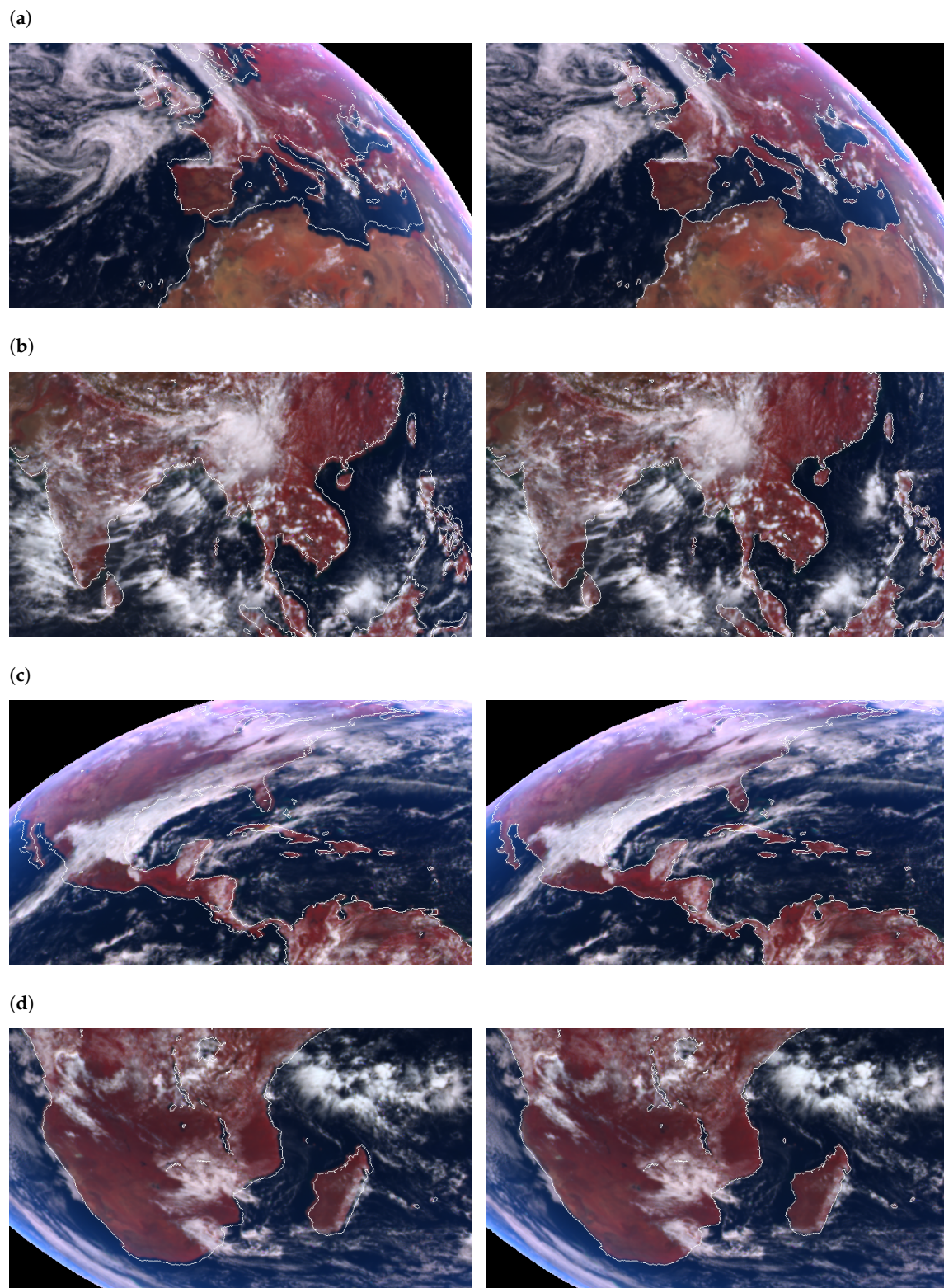
(**a**)



(**b**)



(**c**)



(**d**)



**Figure 7.** Examples of different reflectance images from EPIC L1B version 2 before (**left**) and after (**right**) applying the new registration: (**a**) Europe and North Africa on 4 August 2015 14:37:27 UTC; (**b**) East Asia on 24 July 2016 07:09:19 UTC; (**c**) Caribbean Sea and Gulf of Mexico on 15 December 2017 17:17:03 UTC; and (**d**) the south of Africa and Madagascar on 16 May 2018 09:10:54 UTC.
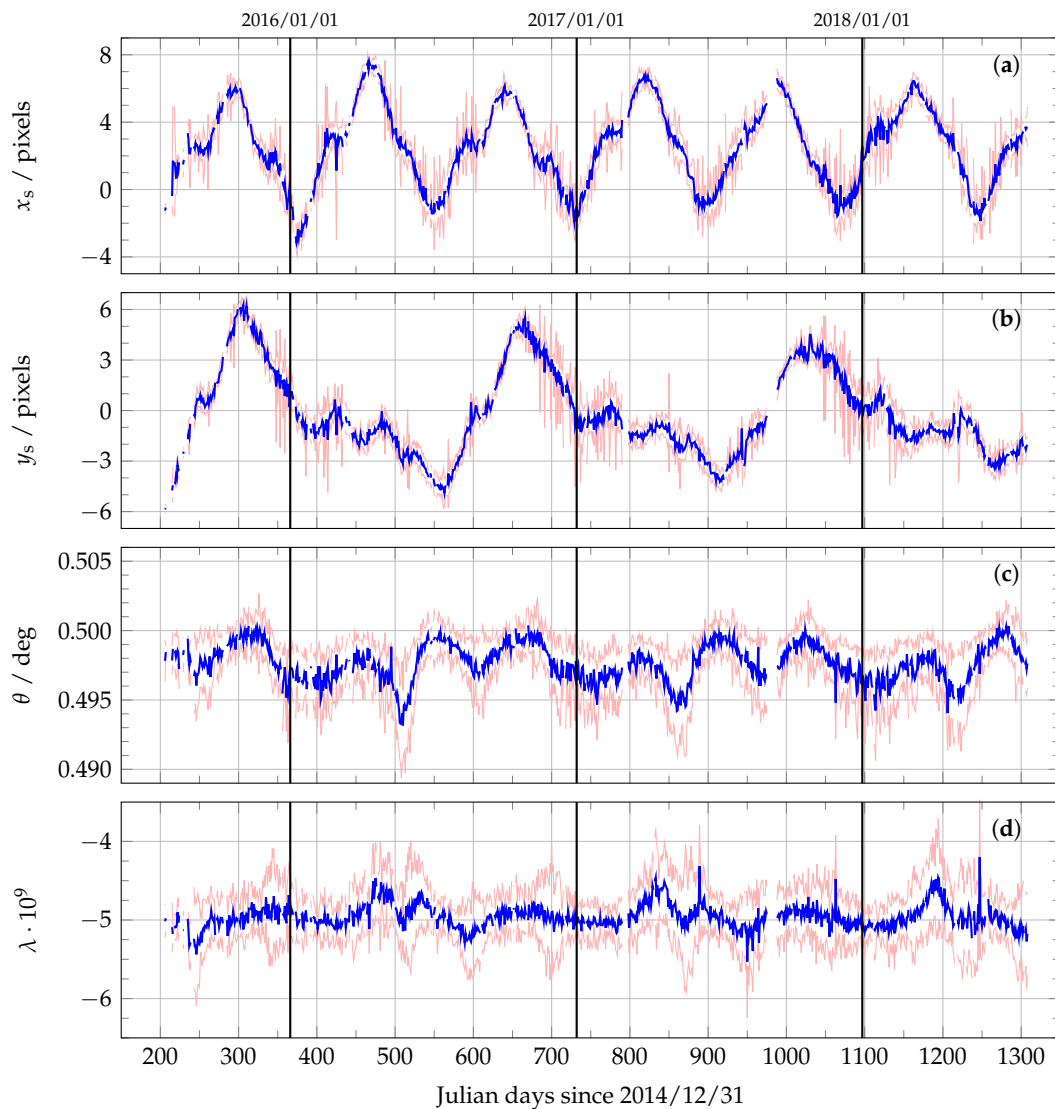
**Figure 8.** Time series plot for every transformation parameter: (**a**) the horizontal shift $x_s$; (**b**) the vertical shift $y_s$; (**c**) the rotation angle $\theta$; and (**d**) the radial distortion parameter $\lambda$. The blue lines represent the evolution of the daily mean value for every parameter, while the light red curves represent the daily dispersion of each parameter given as the daily mean value plus/minus the daily standard deviation.

The time series in Figure 8 can be also interpreted as noisy signals with unevenly sampled data due to the days where there are not enough images to be processed. To estimate the dominant frequencies for every time series, we computed their periodograms by means of the Lomb–Scargle method [40–42]. Before this computation, the signals need to be adapted to have zero mean (subtracting their mean values) and standard deviation of unity (dividing by their standard deviations). In Figure 9, we show the resulting normalised periodogram amplitudes for every transformation parameter as a function of the time period. It can be noticed that three parameters ($x_s$, $y_s$, and $\lambda$) share one common frequency whose period $T_1$ is located in the interval of $[173, 180]$ days, and two parameters ($y_s$ and $\lambda$) share a second frequency whose period $T_2$ is located in the interval of $[363, 366]$ days. The first period $T_1$ corresponds to the 180-day orbit period of DSCOVR's spacecraft around the Lagrangian point $L_1$ [43]. The rotation angle $\theta$ shows different dominant frequencies; however, we do not consider them relevant because the behaviour of its time series is indeed flat around a constant value.
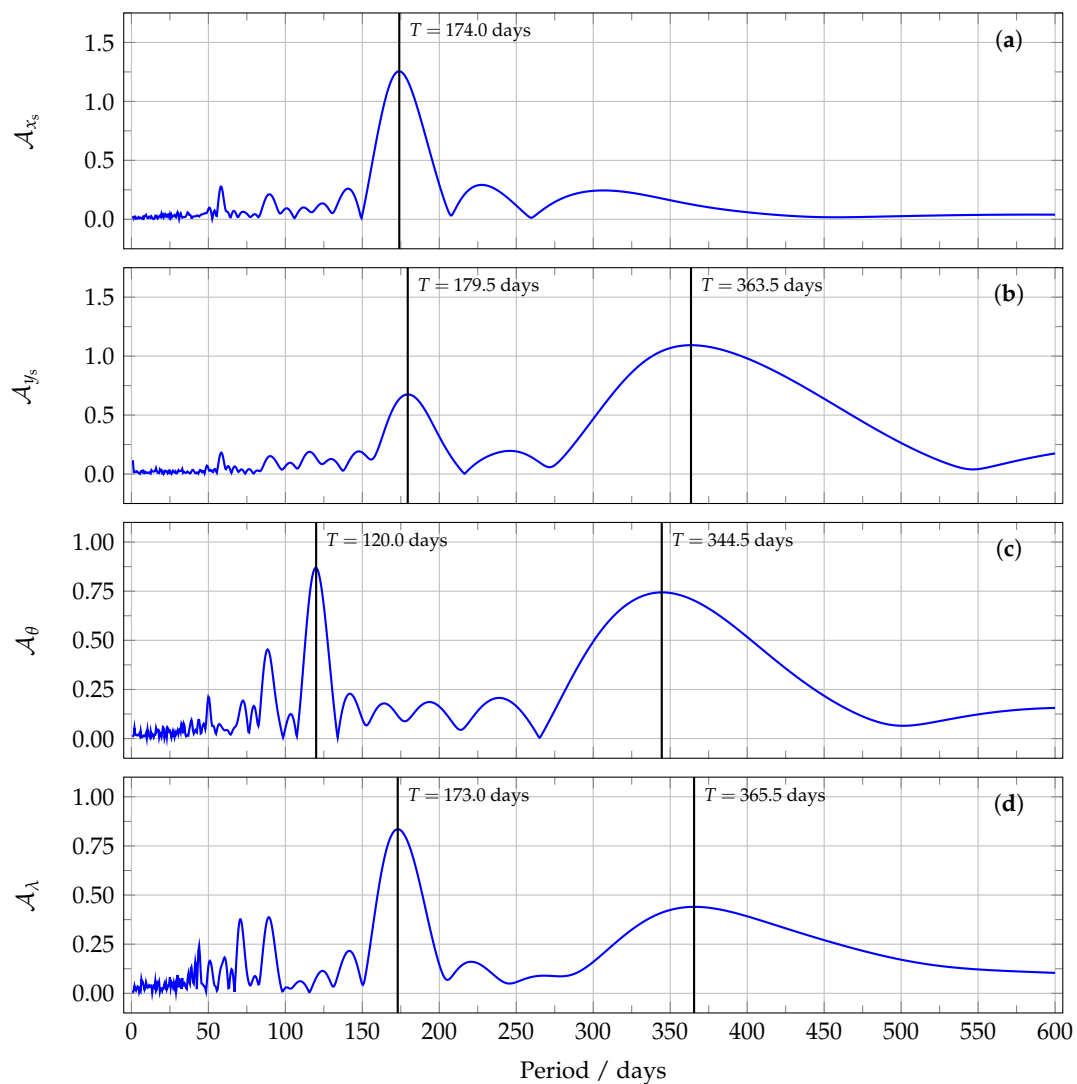
**Figure 9.** Normalised periodogram amplitude $\mathcal{A}$ as a function of the time period in days for every transformation parameter: (**a**) the horizontal shift $x_s$; (**b**) the vertical shift $y_s$; (**c**) the rotation angle $\theta$; and (**d**) the radial distortion parameter $\lambda$.

## 4. Conclusions

In this study, we addressed the geolocation issue that is present in the imagery from EPIC L1B version 2, and to solve it we developed a fast automatic image registration scheme consisting of a motion plus a radial distortion correction. The optimal transformation parameters were computed through Tikhonov least squares minimisation, in which the reference pairs of distorted and registered pixel coordinates were determined automatically by matching features from the theoretical and radiometric coastlines.

We showed that the shift parameters $[x_s, y_s]^\mathsf{T}$ present an oscillatory behaviour whose period is close to the period of DSCOVR's Lissajous orbit, while the rotation angle $\theta$ and the distortion parameter $\lambda$ were found to be more stable and close to the a priori solution that we provided. After the proposed registration, the spatial distances between the common features found in the theoretical and radiometric coastlines are concentrated around the interval of $[1.25, 1.50]$ pixels, a considerable improvement compared to the interval of $[3, 7]$ pixels from the original L1B images. Therefore, we conclude that this procedure enhances the registration of the images from EPIC L1B version 2 with their corresponding arrays of Earth coordinates. This enhancement will have a positive impact in climate science applications from EPIC measurements, as it will reduce the errors in the

retrieval of atmospheric properties due to the inappropriate selection of the surface properties of every pixel. The algorithm developed in this work could be adapted for improving the registration of the future geostationary UVN (Ultraviolet-Visible-Near-infrared) sensors Sentinel-4, GEMS and TEMPO.

This study was based on the OpenCV library, which contains other matching algorithms apart from the brute-force matching of corresponding keypoints, including template matching (e.g., Chamfer matching [44]). In principle, it is interesting to compare their efficiencies for solving the registration problem considered in this article. In addition, the effect of using different robust estimators (e.g., RANSAC) can be analysed in more detail. These topics will be investigated in our future research.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| BRIEF | Binary Robust Independent Elementary Features |
| CCD | Charge-Coupled Device |
| DSCOVR | Deep Space Climate Observatory |
| EPIC | Earth Polychromatic Imaging Camera |
| FAST | Features from Accelerated Segment Test |
| GEMS | Geostationary Environment Monitoring Spectrometer |
| GSHHG | Global Self-consistent Hierarchical, High-resolution Geography database |
| L1B | Level 1B |
| MERIS | MEdium Resolution Imaging Spectrometer |
| NASA | National Aeronautics and Space Administration |
| ORB | Oriented FAST and Rotated BRIEF |
| RANSAC | RANdom SAmple Consensus |
| RGB | Red Green Blue |
| SIFT | Scale-Invariant Feature Transform |
| SURF | Speeded-Up Robust Features |
| TEMPO | Tropospheric Emissions: Monitoring of POllution |
| UVN | Ultraviolet-Visible-Near-infrared |

## References

1. Marshak, A.; Herman, J.; Szabo, A.; Blank, K.; Carn, S.; Cede, A.; Geogdzhayev, I.; Huang, D.; Huang, L.K.; Knyazikhin, Y.; et al. Earth Observations from DSCOVR EPIC Instrument. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 1829–1850. [CrossRef] [PubMed]

2. Marshak, A.; Knyazikhin, Y. The spectral invariant approximation within canopy radiative transfer to support the use of the EPIC/DSCOVR oxygen B-band for monitoring vegetation. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *191*, 7–12. [CrossRef] [PubMed]

3. Herman, J.; Huang, L.; McPeters, R.; Ziemke, J.; Cede, A.; Blank, K. Synoptic ozone, cloud reflectivity, and erythemal irradiance from sunrise to sunset for the whole earth as viewed by the DSCOVR spacecraft from the earth–sun Lagrange 1 orbit. *Atmos. Meas. Tech.* **2018**, *11*, 177–194. [CrossRef]

4. Carn, S.A.; Krotkov, N.A.; Fisher, B.L.; Li, C.; Prata, A.J. First observations of volcanic eruption clouds from the L1 Earth-Sun Lagrange point by DSCOVR/EPIC. *Geophys. Res. Lett.* **2018**, *45*, 11456–11464. [CrossRef]

5. Efremenko, D.S.; Loyola R., D.G.; Hedelt, P.; Spurr, R.J.D. Volcanic SO$_2$ plume height retrieval from UV sensors using a full-physics inverse learning machine algorithm. *Int. J. Remote Sens.* **2017**, *38*, 1–27. [CrossRef]

6. Xu, X.; Wang, J.; Wang, Y.; Zeng, J.; Torres, O.; Yang, Y.; Marshak, A.; Reid, J.; Miller, S. Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 point. *Geophys. Res. Lett.* **2017**, *44*, 7544–7554. [CrossRef]

7. Natraj, V.; Jiang, J.H.; Doicu, A.; Loyola, D.; Kopparla, P.; Yung, Y.L. Aerosol retrievals from DSCOVR measurements. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6026–6028. [CrossRef]

8. Yang, Y.; Meyer, K.; Wind, G.; Zhou, Y.; Marshak, A.; Platnick, S.; Min, Q.; Davis, A.B.; Joiner, J.; Vasilkov, A.; et al. Cloud products from the Earth Polychromatic Imaging Camera (EPIC): Algorithms and initial evaluation. *Atmos. Meas. Tech.* **2019**, *12*, 2019–2031. [CrossRef]

9. Davis, A.B.; Merlin, G.; Cornet, C.; Labonnote, L.C.; Riédi, J.; Ferlay, N.; Dubuisson, P.; Min, Q.; Yang, Y.; Marshak, A. Cloud information content in EPIC/DSCOVR's oxygen A- and B-band channels: An optimal estimation approach. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *216*, 6–16. [CrossRef]

10. Davis, A.B.; Ferlay, N.; Libois, Q.; Marshak, A.; Yang, Y.; Min, Q. Cloud information content in EPIC/DSCOVR's oxygen A- and B-band channels: A physics-based approach. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *220*, 84–96. [CrossRef]

11. Gao, M.; Zhai, P.W.; Yang, Y.; Hu, Y. Cloud remote sensing with EPIC/DSCOVR observations: A sensitivity study with radiative transfer simulations. *J. Quant. Spectrosc. Radiat. Transf.* **2019**, *230*, 56–60. [CrossRef]

12. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]

13. Tikhonov, A.N.; Arsenin, V.I. *Solutions of Ill-Posed Problems*; Scripta Series in Mathematics; Winston & Sons: Washington, DC, USA, 1977.

14. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 228–240. [CrossRef]

15. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Linearized radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 241–251. [CrossRef]

16. Efremenko, D.; Doicu, A.; Loyola, D.; Trautmann, T. Acceleration techniques for the discrete ordinate method. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *114*, 73–81. [CrossRef]

17. Efremenko, D.S.; Doicu, A.; Loyola, D.; Trautmann, T. Fast stochastic radiative transfer models for trace gas and cloud property retrievals under cloudy conditions. In *Springer Series in Light Scattering*; Springer International Publishing: Basel, Switzerland, 2018; pp. 231–277. [CrossRef]

18. del Águila, A.; Efremenko, D.S.; Molina García, V.; Xu, J. Analysis of two dimensionality reduction techniques for fast simulation of the spectral radiances in the Hartley-Huggins band. *Atmosphere* **2019**, *10*, 142. [CrossRef]

19. Blank, K. EPIC Geolocation and Color Imagery Algorithm Revision 5. Technical Report. 2017. Available online: https://eosweb.larc.nasa.gov/project/dscovr/DSCOVR_EPIC_Geolocation_V02.pdf (accessed on 23 July 2019).

20. Haney, C.; Doelling, D.; Minnis, P.; Bhatt, R.; Scarino, B.; Gopalan, A. The calibration of the DSCOVR EPIC multiple visible channel instrument using MODIS and VIIRS as a reference. In Proceedings of the Earth Observing Systems XXI, San Diego, CA, USA, 28 August–1 September 2016. [CrossRef]

21. Available online: https://search.earthdata.nasa.gov (accessed on 27 May 2019).

22. Geogdzhayev, I.V.; Marshak, A. Calibration of the DSCOVR EPIC visible and NIR channels using MODIS Terra and Aqua data and EPIC lunar observations. *Atmos. Meas. Tech.* **2018**, *11*, 359–368. [CrossRef]

23. Fitzgibbon, A.W. Simultaneous linear estimation of multiple view geometry and lens distortion. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; pp. 125–132. [CrossRef]

24. Wang, A.; Qiu, T.; Shao, L. A simple method of radial distortion correction with centre of distortion estimation. *J. Math. Imaging Vis.* **2009**, *35*, 165–172. [CrossRef]

25. Doicu, A.; Trautmann, T.; Schreier, F. *Numerical Regularization for Atmospheric Inverse Problems*; Springer: Berlin/Heidelberg, Germany, 2010. [CrossRef]

26. Wessel, P.; Smith, W.H.F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res. Solid Earth* **1996**, *101*, 8741–8743. [CrossRef]

27. Serra, J. Introduction to mathematical morphology. *Comput. Vis. Graph. Image Process.* **1986**, *35*, 283–305. [CrossRef]

28. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *8*, 679–698. [CrossRef]

29. Available online: https://www.pyimagesearch.com/2015/04/06/zero-parameter-automatic-canny-edge-detection-with-python-and-opencv (accessed on 11 July 2019).

30. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *Photogramm. Rec.* **2014**, *29*, 144–166. [CrossRef]

31. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

32. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]

33. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571. [CrossRef]

34. Kaehler, A.; Bradski, G. *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*; O'Reilly Media: Sebastopol, CA, USA, 2016.

35. Rosten, E.; Porter, R.; Drummond, T. Faster and better: A machine learning approach to corner detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 105–119. [CrossRef] [PubMed]

36. Calonder, M.; Lepetit, V.; Ozuysal, M.; Trzcinski, T.; Strecha, C.; Fua, P. BRIEF: Computing a local binary descriptor very fast. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 1281–1298. [CrossRef]

37. Available online: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_matcher/py_matcher.html (accessed on 11 July 2019).

38. Hamming, R. Error detecting and error correcting codes. *Bell Syst. Tech. J.* **1950**, *29*, 147–160. [CrossRef]

39. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]

40. Lomb, N. Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci.* **1976**, *39*, 447–462. [CrossRef]

41. Scargle, J. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* **1982**, *263*, 835–853. [CrossRef]

42. Hocke, K. Phase estimation with the Lomb-Scargle periodogram method. *Annal. Geophys.* **1998**, *16*, 356–358.

43. Jiang, J.H.; Zhai, A.J.; Herman, J.; Zhai, C.; Hu, R.; Su, H.; Natraj, V.; Li, J.; Xu, F.; Yung, Y.L. Using Deep Space Climate Observatory measurements to study the Earth as an exoplanet. *Astron. J.* **2018**, *156*, 26. [CrossRef]

44. Van Herk, M. Image registration using Chamfer matching. In *Handbook of Medical Image Processing and Analysis*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 591–603. [CrossRef]